

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 02:14:44 ; Search time 4048 Seconds

(without alignments)
16296.467 Million cell updates/sec

Title: US-09-967-237-1

Perfect score: 1522

Sequence: 1 acagtcagccgcatgctcc.....tcaataataattataat 1522

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

- 1: gb_da:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_cm:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_da:*
- 16: em_fun:*
- 17: em_in:*
- 18: em_mu:*
- 19: em_mu:*
- 20: em_mu:*
- 21: em_or:*
- 22: em_or:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hcg_hum:*
- 31: em_hcg_inv:*
- 32: em_hcg_other:*
- 33: em_hcg_mus:*
- 34: em_hcg_pin:*
- 35: em_hcg_rnd:*
- 36: em_hcg_rnd:*
- 37: em_hcg_rnd:*
- 38: em_hcg_vtl:*
- 39: em_hcg_hum:*
- 40: em_hcg_mus:*
- 41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1522	100.0	1522	AR074439	AR074439 Sequence
2	1522	100.0	1522	AR081119	AR081119 Sequence
3	1522	100.0	1522	AR085316	AR085316 Sequence
4	1522	100.0	1522	AR088064	AR088064 Sequence
5	1522	100.0	1522	AR104223	AR104223 Sequence
6	1522	100.0	1522	AR143487	AR143487 Sequence
7	1522	100.0	1522	AR171392	AR171392 Sequence
8	1522	100.0	1522	AR171563	AR171563 Sequence
9	1522	100.0	1522	BD243152	BD243152 MN gene a
10	1522	100.0	1522	AX330007	AX330007 Sequence
11	1522	100.0	1522	AX332607	AX332607 Sequence
12	1522	100.0	1522	AX333244	AX333244 Sequence
13	1522	100.0	1522	AX336174	AX336174 Sequence
14	1522	100.0	1522	HSNATUMN	X66839 H. sapiens M
15	1520.4	99.9	1639	BC014950	BC014950 Homo sapi
16	1517.4	99.7	1519	HAI10588	A010588 Homo sapi
17	1399	91.9	1399	AR095263	AR095263 Sequence
18	759.2	49.9	1665	MMU245857	A0245857 Mus muscu
19	457.8	30.1	1671	AB086322	AB086322 Mus muscu
20	414.4	27.2	6521	AX795690	AX795690 Sequence
21	414.4	27.2	6521	AX822141	AX822141 Sequence
22	414.4	27.2	6521	AX825781	AX825781 Sequence
23	414.4	27.2	10898	AR074442	AR074442 Sequence
24	414.4	27.2	10898	AR081122	AR081122 Sequence
25	414.4	27.2	10898	AR085319	AR085319 Sequence
26	414.4	27.2	10898	AR088067	AR088067 Sequence
27	414.4	27.2	10898	AR104226	AR104226 Sequence
28	414.4	27.2	10898	AR143490	AR143490 Sequence
29	414.4	27.2	10898	AR171395	AR171395 Sequence
30	414.4	27.2	10898	AR171566	AR171566 Sequence
31	414.4	27.2	10898	BD243155	BD243155 MN gene a
32	414.4	27.2	187856	AL157674	AL157674 Human DNA
33	414.4	27.2	415	AF334629	AF334629 Homo sapi
34	413.4	27.2	415	AR074454	AR074454 Sequence
35	413.4	27.2	415	AR081134	AR081134 Sequence
36	413.4	27.2	415	AR085331	AR085331 Sequence
37	413.4	27.2	415	AR088079	AR088079 Sequence
38	413.4	27.2	415	AR104238	AR104238 Sequence
39	413.4	27.2	415	AR143502	AR143502 Sequence
40	413.4	27.2	445	AR171406	AR171406 Sequence
41	413.4	27.2	445	AR171577	AR171577 Sequence
42	413.4	25.0	445	BD243167	BD243167 MN gene a
43	380.2	20.0	10898	HSNMC9	Z54349 H. sapiens M
44	303.8	20.0	408	AX198178	AX198178 Sequence
45	303.8	20.0	521	AX197674	AX197674 Sequence

ALIGNMENTS

RESULT 1
LOCUS AR074439 1522 bp DNA linear PAT 28-AUG-2000
DEFINITION Sequence 1 from patent US 5955075.
ACCESSION AR074439
VERSION AR074439.1 GI:10001194
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1522)
AUTHORS Zavada,J., Pastorekova,S. and Pastorek,J.
TITLE Method of inhibiting tumor growth using antibodies to MN protein
JOURNAL Patent: US 5955075-A 1 21-SEP-1999;
FEATURES Location/Qualifiers

QY 121 CCCAGAGTGGCCCCGATGCAAGAGATTCCCTTGGAGAGAGCTCTTCTGGGGA 180
 Db 121 CCCAGAGTGGCCCCGATGCAAGAGATTCCCTTGGAGAGAGCTCTTCTGGGGA 180
 QY 181 GATACCCACTGGGCGAGAGATCTGCCCCAGTGAAGAGATTACCCAGAGAGAGAT 240
 Db 181 GATACCCACTGGGCGAGAGATCTGCCCCAGTGAAGAGATTACCCAGAGAGAGAT 240
 QY 241 CCACCCGGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 300
 Db 241 CCACCCGGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 300
 QY 301 GAATTAAGCCTTAATCAAGAAAGAGAGGCTCCCTGAGATTAAGAGATCTACCT 360
 Db 301 GAATTAAGCCTTAATCAAGAAAGAGAGGCTCCCTGAGATTAAGAGATCTACCT 360
 QY 361 GAGGCTCTGAGAGATCTCAAGAAAGAGAGATTAAGAGAGAGAGAGAGAGAT 420
 Db 361 GAGGCTCTGAGAGATCTCAAGAAAGAGAGATTAAGAGAGAGAGAGAGAGAT 420
 QY 421 GACCAAGTCAATGGCGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
 Db 421 GACCAAGTCAATGGCGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
 QY 481 GCGGCGCGCTTCCAGTCCCGGAGATATCCGCGCGAGAGAGAGAGAGAGAT 540
 Db 481 GCGGCGCGCTTCCAGTCCCGGAGATATCCGCGCGAGAGAGAGAGAGAGAT 540
 QY 541 CTGCGCGCGCTTCCAGTCCCGGAGATATCCGCGCGAGAGAGAGAGAGAT 600
 Db 541 CTGCGCGCGCTTCCAGTCCCGGAGATATCCGCGCGAGAGAGAGAGAGAT 600
 QY 601 AATAATGCGCAAGTGTGCACTGACCTGCTCTGAGATGAGATGAGATGAGAT 660
 Db 601 AATAATGCGCAAGTGTGCACTGACCTGCTCTGAGATGAGATGAGATGAGAT 660
 QY 661 GGGCGGAGTACCGGCGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 720
 Db 661 GGGCGGAGTACCGGCGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 720
 QY 721 TCGAGACACACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 780
 Db 721 TCGAGACACACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 780
 QY 781 ACCGCTTGGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 840
 Db 781 ACCGCTTGGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 840
 QY 841 GCGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 900
 Db 841 GCGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 900
 QY 901 GAAGAAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 960
 Db 901 GAAGAAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 960
 QY 961 CTGCGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1020
 Db 961 CTGCGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1020
 QY 1021 GCGCGAGAGTGTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1080
 Db 1021 GCGCGAGAGTGTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1080
 QY 1081 CACACCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1140
 Db 1081 CACACCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1140
 QY 1141 GCGAT 1200
 Db 1141 GCGAT 1200
 QY 1201 AGTCTCTGGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1260

Db 1201 AGTCTCTGGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1260
 QY 1261 GCGCTGATTTTGGGCTCTCTTGTGCTGACAGAGAGAGAGAGAGAGAT 1320
 Db 1261 GCGCTGATTTTGGGCTCTCTTGTGCTGACAGAGAGAGAGAGAGAGAT 1320
 QY 1321 AGGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380
 Db 1321 AGGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380
 QY 1381 ACTGAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1440
 Db 1381 ACTGAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1440
 QY 1441 GAGAGCGGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Db 1441 GAGAGCGGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 1501 TTTTAAATTAATTTTAAAT 1522
 Db 1501 TTTTAAATTAATTTTAAAT 1522
 RESULT 3
 LOCUS AR085316 1522 bp DNA linear PAT 01-SEP-2000
 DEFINITION Sequence 1 from patent US 5981711.
 ACCESSION AR085316
 VERSION AR085316.1 GI:10012085
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1522)
 AUTHORS Zavada, J., Pastorekova, S. and Pastorek, J.
 TITLE MN-specific antibodies and hybridomas
 JOURNAL Patent: US 5981711-A 1 09-NOV-1999;
 FEATURES
 source 1..1522
 /organism="Unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Query Match 100.0%; Score 1522; DB 6; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 3,4e-311;
 Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACAGTCACGCGCATGAGTCCCTGTGCGCCAGCCCTGAGCTGCTGATCCCGGCC 60
 Db 1 ACAGTCACGCGCATGAGTCCCTGTGCGCCAGCCCTGAGCTGCTGATCCCGGCC 60
 QY 61 CTGCTCCAGGCTCTCACTGTGCAACTGCTGCTGCTGCTGCTGCTGCTGCT 120
 Db 61 CTGCTCCAGGCTCTCACTGTGCAACTGCTGCTGCTGCTGCTGCTGCTGCT 120
 QY 121 CCCAGAGAGTGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 180
 Db 121 CCCAGAGAGTGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 180
 QY 181 GATGACCCACTGGGCGAGAGATCTGCCCCAGTGAAGAGATTACCCAGAGAGAGAT 240
 Db 181 GATGACCCACTGGGCGAGAGATCTGCCCCAGTGAAGAGATTACCCAGAGAGAGAT 240
 QY 241 CCACCCGGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 300
 Db 241 CCACCCGGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 300
 QY 301 GAATTAAGCCTTAATCAAGAAAGAGAGAGAGAGAGAGAGAGAGAT 360
 Db 301 GAATTAAGCCTTAATCAAGAAAGAGAGAGAGAGAGAGAGAGAGAT 360
 QY 361 GAGGCTCTGAGAGATCTCAAGAAAGAGAGAGAGAGAGAGAGAGAGAT 420

Db 361 GAGGCTCTGGAATCTCTCAAGAACCCCAAGTAATGCCACAGGGAAGAAAGGGAT 420
Qy 421 GACCAAGATCATTTGGGCTATGAGGCGACCCGCTGACCCCGGTTGCCCAAGCTGC 480
Db 421 GACCAAGATCATTTGGGCTATGAGGCGACCCGCTGACCCCGGTTGCCCAAGCTGC 480
Qy 481 GCGGCGCTTCCAGTCCCGGCTGATATCCGCCCCCAAGCTGCGGCTTGTGCCCGGC 540
Db 481 GCGGCGCTTCCAGTCCCGGCTGATATCCGCCCCCAAGCTGCGGCTTGTGCCCGGC 540
Qy 541 CTGCGCCCTTGAATCTCTGAGGCTTCCAGCTCCGCGCTCCAGAACTGCGCTGCG 600
Db 541 CTGCGCCCTTGAATCTCTGAGGCTTCCAGCTCCGCGCTCCAGAACTGCGCTGCG 600
Qy 601 AACATGCGCACAGTGTGACATGACCCGCTCTGAGGCTAGAGATGGCTCTGGTCCC 660
Db 601 AACATGCGCACAGTGTGACATGACCCGCTCTGAGGCTAGAGATGGCTCTGGTCCC 660
Qy 661 GGGCGGAGTACCGGCTCTGAGCTGCACTGAGGCGCTGAGGCTGAGGCTGCGGCG 720
Db 661 GGGCGGAGTACCGGCTCTGAGCTGCACTGAGGCGCTGAGGCTGAGGCTGCGGCG 720
Qy 721 TGGAGACACATGTGGAAGGCAACGCTTCCCTGCGGAGATCCACGTGTTCACTCAGC 780
Db 721 TGGAGACACATGTGGAAGGCAACGCTTCCCTGCGGAGATCCACGTGTTCACTCAGC 780
Qy 781 ACCGCTTTGCGAGATTGACAGAGCTTGGGCGCCCGGAGGCTGCGCGGTTGGCC 840
Db 781 ACCGCTTTGCGAGATTGACAGAGCTTGGGCGCCCGGAGGCTGCGCGGTTGGCC 840
Qy 841 GCTTTCTGAGAGAGGCGCCGGAAGAAACAGTGCCTATGAGCACTTGTCTGCTTGG 900
Db 841 GCTTTCTGAGAGAGGCGCCGGAAGAAACAGTGCCTATGAGCACTTGTCTGCTTGG 900
Qy 901 GAAGAAATGCTGAGAGAGGCTCAGAGCTCAGGCTCCAGAGCTGACATATGCACTC 960
Db 901 GAAGAAATGCTGAGAGAGGCTCAGAGCTCAGGCTCCAGAGCTGACATATGCACTC 960
Qy 961 CTGCTCTGATCTTACGCGCTACTTCCAAATATGAGGAGTCTGACTACAGCGCTGT 1020
Db 961 CTGCTCTGATCTTACGCGCTACTTCCAAATATGAGGAGTCTGACTACAGCGCTGT 1020
Qy 1021 GCCCAGGCTGATCTGAGTGTGTTTAAACAGACAGATGATGCTGAGTGAAGCTC 1080
Db 1021 GCCCAGGCTGATCTGAGTGTGTTTAAACAGACAGATGATGCTGAGTGAAGCTC 1080
Qy 1081 CACACCTCTCTGACACCTCTGAGGACCTGAGTCTCTGCTACAGCTGAACCTTCCGA 1140
Db 1081 CACACCTCTCTGACACCTCTGAGGACCTGAGTCTCTGCTACAGCTGAACCTTCCGA 1140
Qy 1141 GCGAGCGACCTTTGAATGAGGCGAGTGAAGGCTCTTCCCTGCTGAGTGAAGCAAC 1200
Db 1141 GCGAGCGACCTTTGAATGAGGCGAGTGAAGGCTCTTCCCTGCTGAGTGAAGCAAC 1200
Qy 1201 AGTCTCTGAGGCTGAGCGAGTCACTGATATCTGCTGCTGCTGCTGATCACTTA 1260
Db 1201 AGTCTCTGAGGCTGAGCGAGTCACTGATATCTGCTGCTGCTGCTGATCACTTA 1260
Qy 1261 GCGCTGCTTGTGCTCTTCTTGTGCTGACCAAGGCTGCGCTTCTTGTGAGATGAGA 1320
Db 1261 GCGCTGCTTGTGCTCTTCTTGTGCTGACCAAGGCTGCGCTTCTTGTGAGATGAGA 1320
Qy 1321 AGGCAAGCAGAGAGGGAACCAAGAGGAGTGTGAGCTACCGCCACAGAGGTAGCGAG 1380
Db 1321 AGGCAAGCAGAGAGGGAACCAAGAGGAGTGTGAGCTACCGCCACAGAGGTAGCGAG 1380
Qy 1381 ACTGAGGCTTGAAGGCTGAGTCTTGAAGATGAGAGAGCAGAGGCTTGAAGG 1440
Db 1381 ACTGAGGCTTGAAGGCTGAGTCTTGAAGATGAGAGAGCAGAGGCTTGAAGG 1440
Qy 1441 GGAAGCGATACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500

Db 1441 GGAAGCGATACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy 1501 TTTTAAATAAATAATTTTAAAT 1522
Db 1501 TTTTAAATAAATAATTTTAAAT 1522
RESULT 4
AR088064 1522 bp DNA linear PAT 07-SEP-2000
LOCUS AR088064
DEFINITION Sequence 1 from patent US 5989838.
ACCESSION AR088064
VERSION AR088064.1 GI:10014827
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1522)
AUTHORS Zavada, J., Pastorekova, S. and Pastorek, J.
TITLE Immunological methods of detecting MN proteins and MN polypeptides
JOURNAL Patent: US 589838-A 1 23-NOV-1999;
FEATURES
Location/Qualifiers
1..1522
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1522; DB 6; Length 1522;
Best Local Similarity 100.0%; Pred. No. 3, 4e-311;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACACTCAGCCGATAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 1 ACACTCAGCCGATAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Qy 61 CTTGCTCAGGCTCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 CTTGCTCAGGCTCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 121 CCCCAGAGTGTGCCCCGAGTGCAGAGAGATTCCTCCCTGAGAGAGGCTCTTCTGGGAGA 180
Db 121 CCCCAGAGTGTGCCCCGAGTGCAGAGAGATTCCTCCCTGAGAGAGGCTCTTCTGGGAGA 180
Qy 181 GATACCCACTGCGCGAGAGAGATCTGCCAAGTGAAGAGATTCACCCAGAGAGAGAT 240
Db 181 GATACCCACTGCGCGAGAGAGATCTGCCAAGTGAAGAGATTCACCCAGAGAGAGAT 240
Qy 241 CCAACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATCTACCT 300
Db 241 CCAACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATCTACCT 300
Qy 301 GAAGTTAAGCTTAATCAGAAAGAGGCTCTCTGAAGTTAGAGATCTTACTTCTT 360
Db 301 GAAGTTAAGCTTAATCAGAAAGAGGCTCTCTGAAGTTAGAGATCTTACTTCTT 360
Qy 361 GAGGCTCTGAGATCTTCAAGAACCCCAAGTAATGCCCCACAGGGAAGAAAGAGGAT 420
Db 361 GAGGCTCTGAGATCTTCAAGAACCCCAAGTAATGCCCCACAGGGAAGAAAGAGGAT 420
Qy 421 GACCAAGTCAATTTGGGCTATGAGGCGACCCGCTGAGGCTGAGGCTTCCCAAGCTGC 480
Db 421 GACCAAGTCAATTTGGGCTATGAGGCGACCCGCTGAGGCTGAGGCTTCCCAAGCTGC 480
Qy 481 GCGGCGCTTCCAGTCCCGGCTGATATCCGCCCCCAAGCTGCGGCTTGTGCCCGGC 540
Db 481 GCGGCGCTTCCAGTCCCGGCTGATATCCGCCCCCAAGCTGCGGCTTGTGCCCGGC 540
Qy 541 CTGCGCCCTTGAATCTCTGAGGCTTCCAGCTCCGCGCTCCAGAACTGCGCTGCG 600
Db 541 CTGCGCCCTTGAATCTCTGAGGCTTCCAGCTCCGCGCTCCAGAACTGCGCTGCG 600
Qy 601 AACATGCGCACAGTGTGACATGACCCGCTCTGAGGCTAGAGATGGCTCTGGTCCC 660

Dp	60	AAACAATGGCCACAGCTGTGGAACCTGACCCCTGCTCTGGGCTACAGAAATGGCTCTGGGCTCC	660
Qy	661	GGGCGGGAGATACCGGGCTCTGCACTGCACTCTGCACTGGGGGGCTGCAAGTGGTCCGGG	720
Dp	661	GGGCGGGAGATACCGGGCTCTGCACTGCACTCTGCACTGGGGGGCTGCAAGTGGTCCGGG	720
Qy	721	TCGAGACACACTGTGTGAAGGACACGGTTTCCCTGCGAGATCCACAGTGTACCTCAGC	780
Dp	721	TCGAGACACACTGTGTGAAGGACACGGTTTCCCTGCGAGATCCACAGTGTACCTCAGC	780
Qy	781	ACCGCCTTTGCCAGAGTTGACGAGGCTTTGGGGCGCCCGGGAGGCTGGCCGGTGTGGCC	840
Dp	781	ACCGCCTTTGCCAGAGTTGACGAGGCTTTGGGGCGCCCGGGAGGCTGGCCGGTGTGGCC	840
Qy	841	GCCTTTCTGGAGAGGGGCCGGAGAAAAACAGTGGCTATGACAGTTGCTGTCCGCTTG	900
Dp	841	GCCTTTCTGGAGAGGGGCCGGAGAAAAACAGTGGCTATGACAGTTGCTGTCCGCTTG	900
Qy	901	GAAGAAATCGCTGAGAGGCTCAGAGACTCAGGTCCAGGACTGACATATCTGCACCTC	960
Dp	901	GAAGAAATCGCTGAGAGGCTCAGAGACTCAGGTCCAGGACTGACATATCTGCACCTC	960
Qy	961	CTAGCCCTCGACTTACGCGCGTCACTTCCATATAGAGGGGTCTGACTAACCGCCCTGT	1020
Dp	961	CTAGCCCTCGACTTACGCGCGTCACTTCCATATAGAGGGGTCTGACTAACCGCCCTGT	1020
Qy	1021	GCCCAAGGTGTCACTTGATCTGTTTAAACAGACAGTATGCTGATGCTTAAGCAGCTC	1080
Dp	1021	GCCCAAGGTGTCACTTGATCTGTTTAAACAGACAGTATGCTGATGCTTAAGCAGCTC	1080
Qy	1081	CACACCCCTCTGACACCCCTGTGGGGACCTGAGTACTCGGGCTACAGCTGAACCTTCCGA	1144
Dp	1081	CACACCCCTCTGACACCCCTGTGGGGACCTGAGTACTCGGGCTACAGCTGAACCTTCCGA	1144
Qy	1141	GCGACGACAGCCTTTGAATGGCGAGTGAATTGAGGCTCTTCCCTGCTGAGTGAAGCAGC	1200
Dp	1141	GCGACGACAGCCTTTGAATGGCGAGTGAATTGAGGCTCTTCCCTGCTGAGTGAAGCAGC	1200
Qy	1201	AGTCTCTGGGCTGTGAGGCACTCCAGCTGAATTCTGCTGTGCTGTGTGATCATCTTA	1260
Dp	1201	AGTCTCTGGGCTGTGAGGCACTCCAGCTGAATTCTGCTGTGCTGTGTGATCATCTTA	1260
Qy	1261	GCCCTGTTTTTGGCCTCTCTTTTCTGTGACCAAGCGTGGGCTCTTGTGAGATGAGA	1320
Dp	1261	GCCCTGTTTTTGGCCTCTCTTTTCTGTGACCAAGCGTGGGCTCTTGTGAGATGAGA	1320
Qy	1321	AGGCAACAACAAGGGGAAACAAAGGGGCTGTGACTACGCCACAGCAGAGGTAGCCGAG	1380
Dp	1321	AGGCAACAACAAGGGGAAACAAAGGGGCTGTGACTACGCCACAGCAGAGGTAGCCGAG	1380
Qy	1381	ACTGAGAGCTTAGAGGCTGGATCTTGGAGAACTGTGAGAACCCAGCCAGAGGCACTAGGG	1444
Dp	1381	ACTGAGAGCTTAGAGGCTGGATCTTGGAGAACTGTGAGAACCCAGCCAGAGGCACTAGGG	1444
Qy	1441	GAGGCGGTACCTGTCTGTCTGTCTATTAAGCACTTCCCTTTTAACTGCGCAAGAAAT	1500
Dp	1441	GAGGCGGTACCTGTCTGTCTGTCTATTAAGCACTTCCCTTTTAACTGCGCAAGAAAT	1500
Qy	1501	TTTTAAATTAATATTATTAAT	1522
Dp	1501	TTTTAAATTAATATTATTAAT	1522

RESULT 5	AR104223	1522 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	AR104223				
DEFINITION	Sequence 1 from patent US 6093548.				
ACCESSION	AR104223				
VERSION	AR104223.1	GI:12816931			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
	Unclassified.				

```

REFERENCE           1 (bases 1 to 1522)
AUTHORS             Zavada J., Pastorekova S. and Pastorek J.
TITLE               Detection and quantitation of MR-specific antibodies
JOURNAL             Patent: US 6093558-A 1 25-Jul-2000;
FEATURES             Location/Qualifiers
                     1..1522
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

```

```
Query Match      100.0%; Score 1522; DB 6; Length 1522;
Best Local Similarity 100.0%; Pred. No. 3.4e-311;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

QY	1	ACAGTCAGCCGATAGGCTCCCTGTGCCACGACCCCTGGTCTCCCTCTGTGATCCCGGC	60
Db	1	ACAGTAGCCGATAGGCTCCCTGTGCCACGACCCCTGGTCTCCCTCTGTGATCCCGGC	60
QY	61	CTGTCCAGGCTCAGTGTGCACTGTGTCACTGTGTGTGATGCTGTGATCCCGGC	120
Db	61	CTGTCCAGGCTCAGTGTGCACTGTGTGTGATGCTGTGTGATGCTGTGATCCCGGC	120
QY	121	CCCGAAGGTTGCCCCGATGCAAGAGATTTCCCTCTGGAGAGAGCTCTTCTGGGAA	180
Db	121	CCCGAAGGTTGCCCCGATGCAAGAGATTTCCCTCTGGAGAGAGCTCTTCTGGGAA	180
QY	181	GATGACCACTGGGCGAAGAGATCTGCCAGTGAAGAGATTCAACCAGAGAGAGAT	240
Db	181	GATGACCACTGGGCGAAGAGATCTGCCAGTGAAGAGATTCAACCAGAGAGAGAT	240
QY	241	CCACCCGAGAGAGAGATCTACCTGGAGAGAGATCTACCTGGAGAGAGATCTACCT	300
Db	241	CCACCCGAGAGAGAGATCTACCTGGAGAGAGATCTACCTGGAGAGAGATCTACCT	300
QY	301	GAAATTAACTAAATCAGAAAGAGGCTCCCTGAAGTAGAGATCTACTGT	360
Db	301	GAAATTAACTAAATCAGAAAGAGGCTCCCTGAAGTAGAGATCTACTGT	360
QY	361	GAGGCTCCGAGATCTCCAGAACCCCAATTAATGCCACAGGGAACAAAGAGGAT	420
Db	361	GAGGCTCCGAGATCTCCAGAACCCCAATTAATGCCACAGGGAACAAAGAGGAT	420
QY	421	GACCAAGTCAATTGGCGCTATGAGGCGACCCGCTTGACCCCGGTGCCAGCTGC	480
Db	421	GACCAAGTCAATTGGCGCTATGAGGCGACCCGCTTGACCCCGGTGCCAGCTGC	480
QY	481	GGGAGCGGTTCCAGATCCCGAGTGAATCCGCGCCAGCTGCGCGCTCTTGCCGAGC	540
Db	481	GGGAGCGGTTCCAGATCCCGAGTGAATCCGCGCCAGCTGCGCGCTCTTGCCGAGC	540
QY	541	CTGCGCCCTTGGAATCTTGAGGTTCCAGCTCCCGCGCTCCAGAACTGGCGTGGC	600
Db	541	CTGCGCCCTTGGAATCTTGAGGTTCCAGCTCCCGCGCTCCAGAACTGGCGTGGC	600
QY	601	AACAAATGGCACTAGTGTGCACTGACCCCTGCTCTGGAGCTAGAGATGGCTTGAGTCC	660
Db	601	AACAAATGGCACTAGTGTGCACTGACCCCTGCTCTGGAGCTAGAGATGGCTTGAGTCC	660
QY	661	GGGCGGAGTACCGGCTCTGAGCTGATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGC	720
Db	661	GGGCGGAGTACCGGCTCTGAGCTGATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGC	720
QY	721	TGGAGCACTGTGGAAGCAACGTTTCCCTCCGAGATCCAGTGGTTCACTCAGC	780
Db	721	TGGAGCACTGTGGAAGCAACGTTTCCCTCCGAGATCCAGTGGTTCACTCAGC	780
QY	781	ACCGCTTTGACAGATTGACAGAGGCTTGAGGCGCCGAGAGGCTTGCTGTGGCC	840
Db	781	ACCGCTTTGACAGATTGACAGAGGCTTGAGGCGCCGAGAGGCTTGCTGTGGCC	840
QY	841	GCCTTTCTGAGAGAGGCGCGGAAGAAACAGTCTTATAGACAGTGTGTCTGTGCTTG	900
Db	841	GCCTTTCTGAGAGAGGCGCGGAAGAAACAGTCTTATAGACAGTGTGTCTGTGCTTG	900

Dd		1381	ACTGGAGCCTAGAGCTGATCTTGGAAATGTGAAGACCGCAGGCATCTGAGG	1440
Qy		1441	GGAAGCGGTAAGTGCTCTGCTCTCAATAAGCACTGCTTTAATGCCAAGAATT	1500
Dd		1441	GGAAGCGGTAAGTGCTCTGCTCTCAATAAGCACTGCTTTAATGCCAAGAATT	1500
Qy		1501	TTTTAAATTAATATTATAT 1522	
Dd		1501	TTTTAAATTAATATTATAT 1522	
	RESULT 8			
	LOCUS	AR171563	1522 bp DNA	linear PAT 17-DEC-2001
	DEFINITION	Sequence 1 from patent US 6297051.		
	ACCESSION	AR171563		
	VERSION	AR171563.1	GI:17910513	
	KEYWORDS			
	SOURCE	Unknown.		
	ORGANISM	Unknown.		
	REFERENCE	Unclassified.		
	AUTHORS	1 (bases 1 to 1522)		
	TITLE	Zavada,J., Pastorekova,S. and Pastorex,K.J.		
	JOURNAL	MN gene and protein		
	FEATURES	Patent: US 6297051-A 1 02-OCT-2001;		
	SOURCE	Location/Qualifiers		
		1..1522		
		/organism="unknown"		
	ORIGIN	/mol_type="unassigned DNA"		
	Query Match	100.0%; Score 1522; DB 6; Length 1522;		
	Best Local Similarity	100.0%; Pred. No. 3,4e-111;		
	Matches 1522; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy		1 ACAGTCAGCCGCAATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTATCCC GCC	60	
Dd		1 ACAGTCAGCCGCAATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTATCCC GCC	60	
Qy		61 CCTGCTCAGGCTCATGTGTGCACTGCTGTCTACTGTGCTTGATGCTGTGAT	120	
Dd		61 CTTGCTCAGGCTCATGTGTGCACTGCTGTCTACTGTGCTTGATGCTGTGAT	120	
Qy		121 CCCCAAGGTTGCCCGCGATGGAAGAAGATTCCCCCTTGGAGAAGGCTCTTCTGGGAA	180	
Dd		121 CCCCAAGGTTGCCCGCGATGGAAGAAGATTCCCCCTTGGAGAAGGCTCTTCTGGGAA	180	
Qy		181 GATGACCACCTGGGCGAGAGGANTCTGCCAGTGAAGAAGATTCACCGAGAGAGGANT	240	
Dd		181 GATGACCACCTGGGCGAGAGGANTCTGCCAGTGAAGAAGATTCACCGAGAGAGGANT	240	
Qy		241 CCACCCGAGAGAGGATTTACTCTGGAAGAAGATTTACTCTGGAAGAAGATCTACT	300	
Dd		241 CCACCCGAGAGAGGATTTACTCTGGAAGAAGATTTACTCTGGAAGAAGATCTACT	300	
Qy		301 GAAGTTAAGCTTAATCAGAAAGAAGAGGCTCCCTGGAAGTTAAGGATCTAATCTGTT	360	
Dd		301 GAAGTTAAGCTTAATCAGAAAGAAGAGGCTCCCTGGAAGTTAAGGATCTAATCTGTT	360	
Qy		361 GAAGCTCTCTGGAAGATCTCAAGAACCCCAAGAAATTAATCCACAAGGACCAAAGAAGGAT	420	
Dd		361 GAAGCTCTCTGGAAGATCTCTAAGAACCCCAAGAAATTAATCCACAAGGACCAAAGAAGGAT	420	
Qy		421 GACCAGAGTCATTGGCTATGAGAGGCAACCCGCTTGGCCCGGGATGTCCCAAGCTTC	480	
Dd		421 GACCAGAGTCATTGGCTATGAGAGGCAACCCGCTTGGCCCGGGATGTCCCAAGCTTC	480	
Qy		481 GCGGGCGGCTTCAAGTCCCGGTGATATCCGCCCAAGCTCGCGGCTTCTGCGCGGC	540	
Dd		481 GCGGGCGGCTTCAAGTCCCGGTGATATCCGCCCAAGCTCGCGGCTTCTGCGCGGC	540	
Qy		541 CTGAGCCCCCTGGAATCTCTGGGCTTCCAAGTCCCGCGCTCCCAAGACTGAGCTCGCC	600	

Db	541	CTGCGCCCTTGGAACTCTTGAGCTTCCAGCTCCCGCCTTCCCAAACTGCGCTTGGC	600
QY	601	AACAATGGCCACAGTGTGCAACTGACCTGTGCTTGGGCTAGAGATGGCTTGGGTCC	660
Db	601	AACAATGGCCACAGTGTGCAACTGACCTGTGCTTGGGCTAGAGATGGCTTGGGTCC	660
QY	661	GGGGGGGAGATACCGGGCTCTGGAGTGCATGTGCACTGGGGGGCTGACGTCTCGGGC	720
Db	661	GGGGGGGAGATACCGGGCTCTGGAGTGCATGTGCACTGGGGGGCTGACGTCTCGGGC	720
QY	721	TGGAGACAATCTGTGGAAGGCCACCGTTTCCCTCCGAGATCCACGTGTTCACTCAAC	780
Db	721	TGGAGACAATCTGTGGAAGGCCACCGTTTCCCTCCGAGATCCACGTGTTCACTCAAC	780
QY	781	ACGCGCTTTGCCAGAGTTGACGAGGGCTTGGGGGCGCCGGAGAGCGCTGGCCGTGTGGC	840
Db	781	ACGCGCTTTGCCAGAGTTGACGAGGGCTTGGGGGCGCCGGAGAGCGCTGGCCGTGTGGC	840
QY	841	GCCCTTCTGGAAGAGGGCCCGAAGAAAACAGTGCCTATAGACAGTTGCTGTGCTTG	900
Db	841	GCCCTTCTGGAAGAGGGCCCGAAGAAAACAGTGCCTATAGACAGTTGCTGTGCTTG	900
QY	901	GAAGAATGCTGAGGAAGAGCTCAGAGCTCAGGTCCACAGACTGACATATCTGCATC	960
Db	901	GAAGAATGCTGAGGAAGAGCTCAGAGCTCAGGTCCACAGACTGACATATCTGCATC	960
QY	961	CTGCGCTTGAATTCAGCGCGCTACCTCCAAATAGAGGGGTCTGTGACTACACCGCCCT	1020
Db	961	CTGCGCTTGAATTCAGCGCGCTACCTCCAAATAGAGGGGTCTGTGACTACACCGCCCT	1020
QY	1021	GCCGAGAGTGTATCTGACTGTGTGTTAACCAACAGTGAATGCTGAATCTTAAGCACTC	1080
Db	1021	GCCGAGAGTGTATCTGACTGTGTGTTAACCAACAGTGAATGCTGAATCTTAAGCACTC	1080
QY	1081	CACACCTCTCTGACACCTGTGGGGACCTGTGTGACCTGTGAGCTGAACATTCCGA	1140
Db	1081	CACACCTCTCTGACACCTGTGGGGACCTGTGTGACCTGTGAGCTGAACATTCCGA	1140
QY	1141	GCGACGACCTTTGAAATGGCGAGTGAATGAGGCTCTTCCCTGCTGAATGACACG	1200
Db	1141	GCGACGACCTTTGAAATGGCGAGTGAATGAGGCTCTTCCCTGCTGAATGACACG	1200
QY	1201	AGTCTCTGGGGCTGTGAGACCACTGCACTGAATTCCTGCTGGCTCTGTATACATCTTA	1260
Db	1201	AGTCTCTGGGGCTGTGAGACCACTGCACTGAATTCCTGCTGGCTCTGTATACATCTTA	1260
QY	1261	GCCCTGTATTTTGGGCTCTCTTTTGTCTGTACCAAGCGTGCCTTCTGTGCAATGAGA	1320
Db	1261	GCCCTGTATTTTGGGCTCTCTTTTGTCTGTACCAAGCGTGCCTTCTGTGCAATGAGA	1320
QY	1321	AGGACACACAGAAAGGGGAAACMAAGGGGGTGTGAGTACCGCCACACAGATAGCGAG	1380
Db	1321	AGGACACACAGAAAGGGGAAACMAAGGGGGTGTGAGTACCGCCACACAGATAGCGAG	1380
QY	1381	ACTGAGCTTAAGAGCTGATCTTGGAGATGTGAGAGCCAGCCAGAGGCAATCTGAGGG	1440
Db	1381	ACTGAGCTTAAGAGCTGATCTTGGAGATGTGAGAGCCAGCCAGAGGCAATCTGAGGG	1440
QY	1441	GAGAGCGGTAACTGTCTCTGTCTGTATATGCACTTCTCTTTAATCTGCANAATAAT	1500
Db	1441	GAGAGCGGTAACTGTCTCTGTCTGTATATGCACTTCTCTTTAATCTGCANAATAAT	1500
QY	1501	TTTTAAATTAATATTATTAAAT	1522
Db	1501	TTTTAAATTAATATTATTAAAT	1522

```

VERSION      BD243152.1 GI:33052922
KEYWORDS     JP 2002528085-A/1.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        Zavada, J., Pastorekova, S. and Pastorek, J.
JOURNAL      MN gene and protein
              Patent: JP 2002528085-A 1 03-SEP-2002;
              INSTITUTE OF VIROLOGY
COMMENT      OS Homo sapiens (human)
              PN JP 2002528085-A/1
              PD 03-SEP-2002
              PR 22-OCT-1999 JP 2000578465
              PR 23-OCT-1998 US 09/177776, 23-OCT-1998 US 09/178115 PI
              JAN ZAVADA, SILVIA PASTOREKOVA, JAROMIR PASTOREK PC
              C12N15/09, A61K38/00, A61K39/395, A61K48/00, A61P35/00, PC
              007K14/47
              PC C12Q1/02, G01N33/566// (C12Q1/02, C12R1:91), C12N15/00, A61K37/02
              CC MN gene and protein
              FH Key
              FT CDS
              FT mat peptide
              Location/Qualifiers
                1..1522
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"

FEATURES
  source
    1..1522
    Query Match      100.0%; Score 1522; DB 6; Length 1522;
    Best Local Similarity 100.0%; Pred. No. 3,4e-311;
    Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
1  ACAGTCAGCCGCGATGGCTCCCTGTCGCCAGCCGCTCCCTGCTGTGATCCCGGCC
1  ACAGTCAGCCGCGATGGCTCCCTGTCGCCAGCCGCTCCCTGCTGTGATCCCGGCC
60
61  CCTGCTCCAGGCTCCTGTCGCACTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTC
120
61  CCTGCTCCAGGCTCCTGTCGCACTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC
120
121  CCCGAGAGGTTGCCCGGAGTGCAGAGAGATTTCCCTTGGAGAGAGGCTCTTCTGGGAA
180
121  CCCGAGAGGTTGCCCGGAGTGCAGAGAGATTTCCCTTGGAGAGAGGCTCTTCTGGGAA
180
121  CCCGAGAGGTTGCCCGGAGTGCAGAGAGATTTCCCTTGGAGAGAGGCTCTTCTGGGAA
180
181  GATGACCACTGGGCGGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGAT
240
181  GATGACCACTGGGCGGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGAT
240
181  GATGACCACTGGGCGGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGAT
240
181  GATGACCACTGGGCGGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGAT
240
241  CCACCCGGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT
300
241  CCACCCGGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT
300
241  CCACCCGGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT
300
301  GAACTTAAGCTTAATCAAGAAAGAGGGCTCCCTGAAGTTAGAGATTTACTTACTGTT
360
301  GAACTTAAGCTTAATCAAGAAAGAGGGCTCCCTGAAGTTAGAGATTTACTTACTGTT
360
301  GAACTTAAGCTTAATCAAGAAAGAGGGCTCCCTGAAGTTAGAGATTTACTTACTGTT
360
361  GAGGCTCTGAGATCTCTCAAGAACCCAGATTAATGAGGAGAGAGAGAGAGAT
420
361  GAGGCTCTGAGATCTCTCAAGAACCCAGATTAATGAGGAGAGAGAGAGAGAT
420
361  GAGGCTCTGAGATCTCTCAAGAACCCAGATTAATGAGGAGAGAGAGAGAGAT
420
421  GACCAAGATCAITGGGCTATGAGAGGCAACCCGCTGGCCCGGGGTGTCCCAAGCTGC
480
421  GACCAAGATCAITGGGCTATGAGAGGCAACCCGCTGGCCCGGGGTGTCCCAAGCTGC
480
421  GACCAAGATCAITGGGCTATGAGAGGCAACCCGCTGGCCCGGGGTGTCCCAAGCTGC
480
481  GCGGCGGCTTCCAGTCCCGGATGATCCGCGCCAGAGTGGCGGCTTCCGCGGCG
540
481  GCGGCGGCTTCCAGTCCCGGATGATCCGCGCCAGAGTGGCGGCTTCCGCGGCG
540
481  GCGGCGGCTTCCAGTCCCGGATGATCCGCGCCAGAGTGGCGGCTTCCGCGGCG
540
541  CTGGCGCCCTTGAACTCTGGGCTTCCAGCTCCGCGCTCCCAAGTGGCGCTTGGCG
600
541  CTGGCGCCCTTGAACTCTGGGCTTCCAGCTCCGCGCTCCCAAGTGGCGCTTGGCG
600

```

```

DB 541  CTGGCGCCCTTGAACTCTGGGCTTCCAGCTCCGCGCTCCCAAGTGGCGCTTGGCG 600
QY 601  AACATGCGCCAGCTGTGCACTGACCTGCTCTCTGGGCTAGAGATGCTCTGGGCTCC 660
DB 601  AACATGCGCCAGCTGTGCACTGACCTGCTCTCTGGGCTAGAGATGCTCTGGGCTCC 660
QY 661  GGGCGGAGTACCGGGGCTGTGACGCTGATCTGACCTGGGGGCTGAGTGGTCCGGGCT 720
DB 661  GGGCGGAGTACCGGGGCTGTGACGCTGATCTGACCTGGGGGCTGAGTGGTCCGGGCT 720
QY 721  TCGAGACACCTGTGAAAGGCCACCGTTTCCCTGCGGAGATTCACGTGTTCACTCAAC 780
DB 721  TCGAGACACCTGTGAAAGGCCACCGTTTCCCTGCGGAGATTCACGTGTTCACTCAAC 780
QY 781  ACCGCTTTGCGAGAGTACGAGGCTTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGG 840
DB 781  ACCGCTTTGCGAGAGTACGAGGCTTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGG 840
QY 841  GCCTTTGTGAGAGAGGGCCCGGAAACAGTGCCCTATGAGCAGTGTGCTGTGCTGG 900
DB 841  GCCTTTGTGAGAGAGGGCCCGGAAACAGTGCCCTATGAGCAGTGTGCTGTGCTGG 900
QY 901  GAGAAATTCGCTGAGAGAGGCTGAGAGCTGAGAGCTCCGAGATGAGATATCTGACATC 960
DB 901  GAGAAATTCGCTGAGAGAGGCTGAGAGCTGAGAGCTCCGAGATGAGATATCTGACATC 960
QY 961  CTGGCTCTGACTTACGCGCTACTTCAATATGAGGGGCTCTGTGACTACACGCGCTGT 1020
DB 961  CTGGCTCTGACTTACGCGCTACTTCAATATGAGGGGCTCTGTGACTACACGCGCTGT 1020
QY 1021  GCCGAGGCTCATCTGAGATGCTGTTTACAGACAGATATCTGAGTCTAAGACGCTC 1080
DB 1021  GCCGAGGCTCATCTGAGATGCTGTTTACAGACAGATATCTGAGTCTAAGACGCTC 1080
QY 1081  CACAACCTCTCTGACACCTCTGTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081  CACAACCTCTCTGACACCTCTGTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141  GCGACGAGCTTTGATGAGGAGATGATGAGGCTCTCTCTCTGCTGAGTGGACAGC 1200
DB 1141  GCGACGAGCTTTGATGAGGAGATGATGAGGCTCTCTCTCTGCTGAGTGGACAGC 1200
QY 1201  AGTCTGCGGCTCTGAGCAGTCCAGCTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201  AGTCTGCGGCTCTGAGCAGTCCAGCTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261  GCCCTGATTTTGGCTCTCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1261  GCCCTGATTTTGGCTCTCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1321  AGGACAGACAGAAAGGGGAGTGAAGTACCGGCGGAGAGAGAGTGGCGAG 1380
DB 1321  AGGACAGACAGAAAGGGGAGTGAAGTACCGGCGGAGAGAGAGTGGCGAG 1380
QY 1381  ACTGAGACCTTAAGGCTGATTTTGAAGATGTGAAGAACCAAGCAGAGAGATCTGAGG 1440
DB 1381  ACTGAGACCTTAAGGCTGATTTTGAAGATGTGAAGAACCAAGCAGAGAGATCTGAGG 1440
QY 1441  GAGGCGGTAACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 1441  GAGGCGGTAACGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1501  TTTTAAATTAATTTATTAAT 1522
DB 1501  TTTTAAATTAATTTATTAAT 1522

RESULT 10
AX330007 1552 bp DNA linear PAT 09-JAN-2002
LOCUS AX330007
DEFINITION Sequence 516 from Patent WO0194629.
ACCESSION AX330007
VERSION AX330007.1 GI:18102985

```



```

FEATURES
  Source
    Location/Qualifiers
      1..1552
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
  ORIGIN
    Query Match
      100.0%; Score 1522; DB 6; Length 1552;
    Best Local Similarity
      100.0%; Pred. No. 3.4e-311;
    Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 ACAGTCAAGCCGATGAGTCTCCCTGTCAGCCAGCCCTGGCTCCTCTGTGATCCCGGCC 60
  DB 31 ACAGTCAAGCCGATGAGTCTCCCTGTCAGCCAGCCCTGGCTCCTCTGTGATCCCGGCC 90
  QY 61 CCGTCTCCAGGCTCTCACTGTGCAACTGCTGTCTCACTGTCTGTCTGTATGCTGTGCT 120
  DB 91 CCGTCTCCAGGCTCTCACTGTGCAACTGCTGTCTCACTGTCTGTCTGTATGCTGTGCT 150
  QY 121 CCCAGAGGTTGCCCCGAGATGACAGAGATTCCTCCCTTGGAGAGGCTCTTCTGGGAA 180
  DB 151 CCCAGAGGTTGCCCCGAGATGACAGAGATTCCTCCCTTGGAGAGGCTCTTCTGGGAA 210
  QY 181 GATGACCACTGGGCGAGAGAGATCTGCCCAGTGAAGAGATTCAACCAAGAGAGAGAT 240
  DB 211 GATGACCACTGGGCGAGAGAGATCTGCCCAGTGAAGAGATTCAACCAAGAGAGAGAT 270
  QY 241 CCACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATCTACCT 300
  DB 271 CCACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATCTACCT 330
  QY 301 GAAGTTAAGCTTAATCAAGAGAGAGAGGCTCCCTGAAGTTAGAGATCTACCTACTTT 360
  DB 331 GAAGTTAAGCTTAATCAAGAGAGAGAGGCTCCCTGAAGTTAGAGATCTACCTACTTT 390
  QY 361 GAAGGCTCTGAGAGATCTCAAGAACCCAGAAATATGCCCAGAGAGAGAGAGAGAT 420
  DB 391 GAAGGCTCTGAGAGATCTCAAGAACCCAGAAATATGCCCAGAGAGAGAGAGAGAT 450
  QY 421 GACAGAGATCTGAGGCTATGAGAGGAGACCCGCTGAGCCCGGGTGTCCCAAGCTGC 480
  DB 451 GACAGAGATCTGAGGCTATGAGAGGAGACCCGCTGAGCCCGGGTGTCCCAAGCTGC 510
  QY 481 GCGGCGCGCTTCAAGTCCCGGAGATATCCGCCCCAGCTGCGCGCTTCTGCGCGGCC 540
  DB 511 GCGGCGCGCTTCAAGTCCCGGAGATATCCGCCCCAGCTGCGCGCTTCTGCGCGGCC 570
  QY 541 CTGCGCGCCCTGGAATCTCGGAGCTTCAGAGCTCCGCGCTCCAGAACTCGCGCTGCG 600
  DB 571 CTGCGCGCCCTGGAATCTCGGAGCTTCAGAGCTCCGCGCTCCAGAACTCGCGCTGCG 630
  QY 601 AACATGAGCCACAGTGTGCACTGACCTGCTCTCTGGGCTAGAGATGGCTCTGGGTCC 660
  DB 631 AACATGAGCCACAGTGTGCACTGACCTGCTCTCTGGGCTAGAGATGGCTCTGGGTCC 690
  QY 661 GGGCGGAGATGAGGAGCTGAGAGCTGATGCACTGGGAGGAGCTGAGAGCTGCGCGGC 720
  DB 691 GGGCGGAGATGAGGAGCTGAGAGCTGATGCACTGGGAGGAGCTGAGAGCTGCGCGGC 750
  QY 721 TCGAGAGACACTGTGAGAGGACACGTTTCTCTGCGAGATCAAGTGTTCACCTGAGC 780
  DB 751 TCGAGAGACACTGTGAGAGGACACGTTTCTCTGCGAGATCAAGTGTTCACCTGAGC 810
  QY 781 ACCGCTTTTCCAGAGATGACAGAGCTTGGGAGCGCCGAGAGGCTGAGCGGTGTGGCC 840
  DB 811 ACCGCTTTTCCAGAGATGACAGAGCTTGGGAGCGCCGAGAGGCTGAGCGGTGTGGCC 870
  QY 841 GCGTTTCTGAGAGAGGAGCGCGAGAGAAACAGTGCCTATGAGCACTTCTGCTGCTTG 900
  DB 871 GCGTTTCTGAGAGAGGAGCGCGAGAGAAACAGTGCCTATGAGCACTTCTGCTGCTTG 930
  QY 901 GAAGAAATGCTGAGAGAGGCTCAGAGACTCAGGCTCCAGAGCTGGAATATCTGCACTC 960

```

```

  DB 931 GAAGAAATGCTGAGAGAGGCTCAGAGACTCAGGCTCCAGAGACTGAGCATATCTGCACTC 990
  QY 961 CTGCGCTTCTGAGTTCAGCGGCTTCTCAATATGAGAGGAGTCTGATACACCGCCCTGT 1020
  DB 991 CTGCGCTTCTGAGTTCAGCGGCTTCTCAATATGAGAGGAGTCTGATACACCGCCCTGT 1050
  QY 1021 GCGGAGGCTGATCTGAGCTGTGTTTAAACAGAGAGATGCTGAGTCTTAACAGCTC 1080
  DB 1051 GCGGAGGCTGATCTGAGCTGTGTTTAAACAGAGAGATGCTGAGTCTTAACAGCTC 1110
  QY 1081 CACACCTCTGAGACCTCTGTGGAGCTGAGTCACTCTGAGTCAAGCTGAATTCGGA 1140
  DB 1111 CACACCTCTGAGACCTCTGTGGAGCTGAGTCACTCTGAGTCAAGCTGAATTCGGA 1170
  QY 1141 GCGAGCGACCTTTGATAGGAGAGATGATGAGGCTCTCTCTGCTGAGTGAACAGC 1200
  DB 1171 GCGAGCGACCTTTGATAGGAGAGATGATGAGGCTCTCTCTGCTGAGTGAACAGC 1230
  QY 1201 AGTCTCGGGCTGTCAGGCAAGTCAAGTCAATTCCTGCTGCTGCTGAGTGAATCCTA 1260
  DB 1231 AGTCTCGGGCTGTCAGGCAAGTCAAGTCAATTCCTGCTGCTGCTGAGTGAATCCTA 1290
  QY 1261 GCGCTGCTTTTGGGCTCTCTTTTGTCTGTCACAGAGCTGCGCTTCTTGTGCAATGAG 1320
  DB 1291 GCGCTGCTTTTGGGCTCTCTTTTGTCTGTCACAGAGCTGCGCTTCTTGTGCAATGAG 1350
  QY 1321 AGGAGCAGACAGAGAGGAGAACCAAGAGGAGTGTGAGCTACCGCCAGAGAGAGTACGAG 1380
  DB 1351 AGGAGCAGACAGAGAGGAGAACCAAGAGGAGTGTGAGCTACCGCCAGAGAGAGTACGAG 1410
  QY 1381 ACTGAGCTTACAGGCTGAGTCTTGTGAGATGTGAGAGACCAAGAGAGAGATTTGAGGG 1440
  DB 1411 ACTGAGCTTACAGGCTGAGTCTTGTGAGATGTGAGAGACCAAGAGAGAGATTTGAGGG 1470
  QY 1441 GGAGCGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
  DB 1471 GGAGCGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1530
  QY 1501 TTTTAAATTAATTAATTAATTAAT 1522
  DB 1531 TTTTAAATTAATTAATTAATTAAT 1552

  RESULT 12
  AX333244 1552 bp DNA linear PAT 09-JAN-2002
  LOCUS AX333244
  DEFINITION Sequence 3753 from Patent WO0194629.
  ACCESSION AX333244
  VERSION AX333244.1 GI:18123878
  KEYWORDS
  SOURCE
  ORGANISM
    Homo sapiens (human)
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

  REFERENCE
    1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
      Horrigan, S., Soppet, D.R. and Weaver, Z.
      Cancer gene determination and therapeutic screening using signature
      gene sets
      Patent: WO 0194629-A 3753 13-DEC-2001;
      Avallon Pharmaceuticals (US)

  JOURNAL
    Location/Qualifiers
      1..1552
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
  ORIGIN
    Query Match
      100.0%; Score 1522; DB 6; Length 1552;
    Best Local Similarity
      100.0%; Pred. No. 3.4e-311;
    Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1 ACAGTCAAGCCGATGAGTCTCCCTGTCAGCCAGCCCTGGCTCCTCTGTGATCCCGGCC 60

```


Db	31	ACAGTCAGCCGATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTGTGTATCCCCGGC	90
OY	61	CTGTCTCAGGCTTACTGTGCACTGTGTCTGTACTGTGCTTTGATGTGCTTGCAT	120
Db	91	CTGTCTCCAGGCTCCTACTGTCAACTGTGTGTGTCACTGTGCTTTGTATGTGCTTGCAT	150
OY	121	CCCCAGAGTTGCCCGATGAGAGAGATCCCCCTTGGAGAGAGCTCTTCTGGGAA	180
Db	151	CCCCAGAGTTGCCCGATGAGAGAGATTTCCCCCTTGGAGAGAGCTCTTCTGGGAA	210
OY	181	GATGATCCACTGGGCGAGAGAGATTGTGCCAGTGAAGAGATTCACTCCAGAGAGAGAT	240
Db	211	GATGATCCACTGGGCGAGAGAGATTGTGCCAGTGAAGAGATTCACTCCAGAGAGAGAT	270
OY	241	CCAGCCGAGAGAGAGATTACTGTGAGAGAGAGATTACTGTGAGAGAGAGATTACT	300
Db	271	CCAGCCGAGAGAGAGATTACTGTGAGAGAGAGATTACTGTGAGAGAGAGATTACT	330
OY	301	GAAGTTAAGCTTAAATCAGAGAGAGAGGCTCCTGAAGTTAAGAGATTACTTCTGT	360
Db	331	GAAGTTAAGCTTAAATCAGAGAGAGAGGCTCCTGAAGTTAAGAGATTACTTCTGT	390
OY	361	GAGGCTCTGAGATCTCAAGAACCCAGAAATATCCACAGAGGACAAAGAGAGAT	420
Db	391	GAGGCTCTGAGATCTCAAGAACCCAGAAATATCCACAGAGGACAAAGAGAGAT	450
OY	421	GACCAAGATATTGGGCGATATGAGAGGCACTCCGCTTGGCCCGGGTGTCCCACTGC	480
Db	451	GACCAAGATATTGGGCGATATGAGAGGCACTCCGCTTGGCCCGGGTGTCCCACTGC	510
OY	481	GCGGCGGCTTCCAGTCCCCGGTGAATCCGCCCCAGCTGCGCTTGTGCCCGGC	540
Db	511	GCGGCGGCTTCCAGTCCCCGGTGAATCCGCCCCAGCTGCGCTTGTGCCCGGC	570
OY	541	CTGCGCCCCCTGGAATCCTTGGGCTTCCAGTCCCCGCTCTCCAGAACTGGGCTGGG	600
Db	571	CTGCGCCCCCTGGAATCCTTGGGCTTCCAGTCCCCGCTCTCCAGAACTGGGCTGGG	630
OY	601	AAGCAATGACACAGTGTGCACTGAACCTGTGCTCTGGGCTAGAGATGGCTGTGGATCC	660
Db	631	AACAAATGACACAGTGTGCACTGAACCTGTGCTCTGGGCTAGAGATGGCTGTGGATCC	690
OY	661	GAGCGGAGATACCGGCTCTGAGGTGATCTGGAATGGGGGCTGCAAGTGTGCGGGC	720
Db	691	GAGCGGAGATACCGGCTCTGAGGTGATCTGGAATGGGGGCTGCAAGTGTGCGGGC	750
OY	721	TGAGAGCACTGTGAGAGGCCACCGTTTCCCTGCGAGATTCACAGTGTTCACCTCAGC	780
Db	751	TGAGAGCACTGTGAGAGGCCACCGTTTCCCTGCGAGATTCACAGTGTTCACCTCAGC	810
OY	781	AACGCTTTTGCAGATTTGAAGAGCCCTTGGGGGCGCCGGAGGCTGGCGGTGGGC	840
Db	811	AACGCTTTTGCAGATTTGAAGAGCCCTTGGGGGCGCCGGAGGCTGGCGGTGGGC	870
OY	841	GCCTTTCTGAGAGAGGCCCCGAGAACAAACAGTCCATAGACAGTTGTCTGTGCTTG	900
Db	871	GCCTTTCTGAGAGAGGCCCCGAGAACAAACAGTCCATAGACAGTTGTCTGTGCTTG	930
OY	901	GAGGAATGCTGAGAGAGGCTCAGAGACTCAGGCTCCAGACTGGAATATGTGACTC	960
Db	931	GAGGAATGCTGAGAGAGGCTCAGAGACTCAGGCTCCAGACTGGAATATGTGACTC	990
OY	961	CTGCTCTGACTTCAAGCCGCTACTTCCAAATATGAGGGGTCTGACTACCGCCCTGT	1020
Db	991	CTGCTCTGACTTCAAGCCGCTACTTCCAAATATGAGGGGTCTGACTACCGCCCTGT	1050
OY	1021	GCCAGAGGTCTCATCTGACTGTGTTTAAACAGACAGAGATGTCTGATGTAAAGACTC	1080
Db	1051	GCCAGAGGTCTCATCTGACTGTGTTTAAACAGACAGAGATGTCTGATGTAAAGACTC	1110
OY	1081	CACACCTCTTGAACCTGTGGGAGCTGTGTGCTCTGTGGCTACAGCTGAATTTCCGA	1140

Db	111	CACACCCTCTCGACACCCCTGTGGGAGACCTGTGAGCTCTGGCTACACTGAACCTCCGA	1170
Qy	1141	GGGAGCGAGACCTTTGAATGGGCGCAGATATGAGACCCCTCCTTCCGCTGAGAGTGACACG	1200
Ds	1171	GGGAGCGAGACCTTTGAATGGGCGCAGATATGAGACCCCTCCTTCCGCTGAGAGTGACACG	1230
Qy	1201	AGTCTCGGGCTGTGTGAGCCAGTCTCAGCTGAATTCCTGCTGGCTGGCTGTGTGACATCCTA	1260
Ds	1231	AGTCTCGGGCTGTGTGAGCCAGTCTCAGCTGAATTCCTGCTGGCTGGCTGTGTGACATCCTA	1290
Qy	1261	GGCCGTGTTTTTGGCGCTCTTTTGGCTGTACACAGAGTGGGCTCCTGTGTGAGATGAGA	1320
Ds	1291	GGCCGTGTTTTTGGCGCTCTTTTGGCTGTACACAGAGTGGGCTCCTGTGTGAGATGAGA	1350
Qy	1321	AGGCGACACAGAGAGGGAGCCAAAGGGGGTGTGAGCTACCGCCACAGACAGGTACCGAG	1380
Ds	1351	AGGCGACACAGAGAGGGAGCCAAAGGGGGTGTGAGCTACCGCCACAGACAGGTACCGAG	1410
Qy	1381	ACTGAGAGCTTAGAGGCTGGATCTTTGGAGAAATGTGAAGACCGACCGACAGGACATCGAGG	1440
Ds	1411	ACTGAGAGCTTAGAGGCTGGATCTTTGGAGAAATGTGAAGACCGACCGACAGGACATCGAGG	1470
Qy	1441	GGAGCCGGTAACGTCTCTGTCTGTGCTCATTTATSCCATCTTCTTTAAGCCAGAAATT	1500
Ds	1471	GGAGCCGGTAACGTCTCTGTCTGTGCTCATTTATSCCATCTTCTTTAAGCCAGAAATT	1530
Qy	1501	TTTTAAATTAATATTTATTAAT	1522
Ds	1531	TTTTAAATTAATATTTATTAAT	1552
RESULT 13			
AX336174	AX336174	1552 bp	DNA
LOCUS	Sequence 6683 from Patent WO0194629.		linear
DEFINITION	AX336174		
ACCESSION	AX336174.1	GI:18126893	
VERSION			
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,		
JOURNAL	Horrigan, S., Soppet, D.R. and Weaver, Z.		
FEATURES	Cancer gene determination and therapeutic screening using signature		
source	gene sets		
	Patent: WO 0194629-A 6683 13-DEC-2001;		
	Avalon Pharmaceuticals (US)		
	Location/Qualifiers		
	1..1552		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	100.0%; Score 1522;	DB 6;	Length 1552;
Best Local Similarity	100.0%; Pred. No. 3,4e-311;		
Matches 1522;	Conservative 0;	Mismatches 0;	Indels 0;
		Gaps 0;	
Ds	1	ACAGTCAGCCGATGGCTCCCTGTGCCCCAGCCCGTGGCTCCCTGTGTGATCCCGGC	60
Qy	31	ACAGTCAGCCGATGGCTCCCTGTGCCCCAGCCCGTGGCTCCCTGTGTGATCCCGGC	90
Ds	61	CTGTCTCAGGCTCTACATGTGCACTGTGTGTCTACTGTGTTTGTATGGCTCTCAT	120
Qy	91	CTGTCTCAGGCTCTACATGTGCACTGTGTGTCTACTGTGTTTGTATGGCTCTCAT	150
Ds	121	CCCCAGAGATTCCCCCGATCAGAGAGATTCCCTTTGGAGAGAGCTCTTTGGAGNA	180
Qy	151	CCCCAGAGATTCCCCCGATCAGAGAGATTCCCTTTGGAGAGAGCTCTTTGGAGNA	210
Ds	181	GATGACCCACTGGGCGAGAGAGATCTGCCACATGAAAGATTCACCCAGAGAGAGAT	240

Db	211	GATGACCCACTGGGCGAAGGAGTCTGCCCACTGAAGAAGATTCA	CCGAGAAGAGAT	270
Qy	241	CCACCCGGAGAGAGATCTACCTGGAAGAGATCTACCTGGAAGAGATCTACCT		300
Db	271	CCACCCGGAGAGAGATCTACCTGGAAGAGAGATCTACCTGGAAGAGATCTACCT		330
Qy	301	GAAATTAAAGCTAAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGATCTACCTCTGTT		360
Db	331	GAAATTAAAGCTAAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGATCTACCTCTGTT		390
Qy	361	GAGGCTCTTGAGATCTCTAATAAACCCCAATATATCCCAACGGACAAAGAGGAT		420
Db	391	GAGGCTCTTGAGATCTCTAATAAACCCCAATATATCCCAACGGACAAAGAGGAT		450
Qy	421	GACCAGATCATTTGGGCTATGAGGCGACCCGGCTGGGCCGGGATGCCAGCTGC		480
Db	451	GACCAGATCATTTGGGCTATGAGGCGACCCGGCTGGGCCGGGATGCCAGCTGC		510
Qy	481	GCGGGCGGCTTCCAGTCCCGGCTGGATATCCGGCCCAAGCTGCAGCTCTTGCCCGGC		540
Db	511	GCGGGCGGCTTCCAGTCCCGGCTGGATATCCGGCCCAAGCTGCAGCTCTTGCCCGGC		570
Qy	541	CTGCGGCCCTTGAACTCTTGAGCTTCCAGCTCCCGCGCTCCAGAACTGGCTGCAC		600
Db	571	CTGCGGCCCTTGAACTCTTGAGCTTCCAGCTCCCGCGCTCCAGAACTGGCTGCAC		630
Qy	601	AACAATGGCCACAGTGTGGAATGACCCCTGCTCTGGGGCTAAGATATGGCTCTGGATCC		660
Db	631	AACAATGGCCACAGTGTGGAATGACCCCTGCTCTGGGGCTAAGATATGGCTCTGGATCC		690
Qy	661	GGGCGGGAGTACCGGAGCTCTGCACTGTGACCTGCACTGAGGAGGCTCAGATGTCGGAGC		720
Db	691	GGGCGGGAGTACCGGAGCTCTGCACTGTGACCTGCACTGAGGAGGCTCAGATGTCGGAGC		750
Qy	721	TGGAGACAACATGTGGAAGGCCACCGTTCCCTGCGAGATCCAAGTGTACCTCAACCTCAGC		780
Db	751	TGGAGACAACATGTGGAAGGCCACCGTTCCCTGCGAGATCCAAGTGTGTCACTCAACCTCAGC		810
Qy	781	ACCGCTTTGCGAAGTTGACGAGGCTTTGGGAGCGCCCGGAGGCTTGCCGATTTGAGC		840
Db	811	ACCGCTTTGCGAAGTTGACGAGGCTTTGGGAGCGCCCGGAGGCTTGCCGATTTGAGC		870
Qy	841	GCGTTTCTGGAAGAGGCCCGGAGAAACACGTGCTAATAGAGAGTGTGCTGCGCTTG		900
Db	871	GCGTTTCTGGAAGAGGCCCGGAGAAACACGTGCTAATAGAGAGTGTGCTGCTGCGCTTG		930
Qy	901	GAAAGAAATGCTGAGAGAGGCTCAGAGCTCAGGCTCCAGAGCTGACATATCTGCACTC		960
Db	931	GAAAGAAATGCTGAGAGAGGCTCAGAGCTCAGGCTCCAGAGCTGACATATCTGCACTC		990
Qy	961	CTGCGCTCTGACCTTACGCGGCTACTTCCAAATATAGAGGCTCTGACTCAACCGGCTGT		1020
Db	991	CTGCGCTCTGACCTTACGCGGCTACTTCCAAATATAGAGGCTCTGACTCAACCGGCTGT		1050
Qy	1021	GCCCAAGGCTCATCTGACTGTGTTAACCAACAGCAAGTATGCTGAGTCTTAAGCACTC		1080
Db	1051	GCCCAAGGCTCATCTGACTGTGTTAACCAACAGCAAGTATGCTGAGTCTTAAGCACTC		1110
Qy	1081	CACACCTCTCTGACACCCCTGAGGGACCTGATACCTGCGGCTAACAGCTGAACCTCGA		1140
Db	1111	CACACCTCTCTGACACCCCTGAGGGACCTGATACCTGCGGCTAACAGCTGAACCTCGA		1170
Qy	1141	GCGACGACGCTTTGATATGGGAGAGTATGAGGCTCTCTTCCCTGCTGAGTGAACAC		1200
Db	1171	GCGACGACGCTTTGATATGGGAGAGTATGAGGCTCTCTTCCCTGCTGAGTGAACAC		1230
Qy	1201	AGTCTCTGGGCTGTAGGCAAGTCTCAGGTGAATTCCTGCTGGCTGCTGATACATCTCA		1260
Db	1231	AGTCTCTGGGCTGTAGGCAAGTCTCAGGTGAATTCCTGCTGGCTGCTGATACATCTCA		1290
Qy	1261	GCCCTGATTTTGGCTCTTTTGTCTGTACCAAGCTGCGCTTCTGTGAGATGAGA		1320

DB	1291	GGCCGTGTTTGGGCTCCTTTTGGCTGTGCACAGGGTGGCGTTCCTTGTCAGATGAGA	1350
OY	1321	AGCGACGACAGAAAGGGGAACCAAAAGGGGGGTGTAGCTACCGCCACGACAGGTACCCAG	1380
Db	1351	AGGACACACAGAAAGGGGAACCAAAAGGGGGGTGTAGCTACCGCCACGACAGGTACCCAG	1410
OY	1381	ACTGAGCCCTTAAAGGCTTGATCTTGGAGAAATGTGAGAACCCAGCAGACGATCTGAGG	1440
Db	1411	ACTGAGCCCTTAAAGGCTTGATCTTGGAGAAATGTGAGAACCCAGCAGACGATCTGAGG	1470
OY	1441	GAGGCCGGTAACTGTCTCTGTCTCTCTCTATTAATGCCCTTCTTTAACTGCCAAGAAAT	1500
Db	1471	GAGGCCGGTAACTGTCTCTGTCTCTCTCTATTAATGCCCTTCTTTAACTGCCAAGAAAT	1530
OY	1501	TTTTAAATTAATATTTATTAAT	1522
Db	1531	TTTTAAATTAATATTTATTAAT	1552
RESULT 14			
LOCUS	HSMTATNM	1552 bp	RNA linear PRI 10-OCT-1995
DEFINITION	H.sapiens Matu MN mRNA for p54/58n protein.		
ACCESSION	X66839		
VERSION	X66839.1 GI:1000701		
KEYWORDS	transmembrane glycoprotein.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Pastorek, J.		
TITLE	1 (bases 1 to 1552)		
JOURNAL	Direct Submission		
REFERENCE	Submitted (11-JUN-1992) J. Pastorek, Institute of Virology, Slovak Academy of Sciences, Dubravskva 9, 842 46 Bratislava, SLOVAK REPUBLIC		
REMARK	revised by [3] MAT		
AUTHORS	2 (bases 1 to 1552)		
REFERENCE	Pastorek, J., Pastorekova, S., Callebaut, I., Mornon, J., Zelnik, V., Pavarsky, R., Zatoricova, M., Ila, S., Portetelle, D., Standridge, E. J., Zavada, J., and Burny, A.		
TITLE	Cloning and characterization of MN, a human tumor-associated protein with a domain homologous to carbonic anhydrase and a putative helix-loop-helix DNA binding segment		
JOURNAL	Oncogene 9 (10), 2877-2888 (1994)		
MEDLINE	94368734		
REFERENCE	8084592		
AUTHORS	3 (bases 1 to 1552)		
TITLE	Pastorek, J.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (19-JUN-1994) J. Pastorek, Institute of Virology, Slovak Academy of Sciences, Dubravskva 9, 842 46 Bratislava, SLOVAK REPUBLIC		
REMARK	revised by [4] MAT		
REFERENCE	4 (bases 1 to 1552)		
AUTHORS	Pastorek, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-SEP-1995) J. Pastorek, Institute of Virology, Slovak Academy of Sciences, Dubravskva 9, 842 46 Bratislava, SLOVAK REPUBLIC		
COMMENT	On Sep 29, 1995 this sequence version replaced gi:558593.		
FEATURES	location/Qualifiers		
source	1..1552		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..1519		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="MN1"		
	/cell_type="epithelial"		
	/tissue_type="carcinoma"		
source			

[illegible][illegible]

```
QY 1381 ACTGGAGCCTAGAGCCTGATCTTGAGAAATGTGAGAGCCAGCCAGGCAATCTGAGG 1440
DB 1381 ACTGGAGCCTAGAGCCTGATCTTGAGAAATGTGAGAGCCAGCCAGGCAATCTGAGG 1440
QY 1441 GGAGCCGGTAACTGCTCTGCTCTGCTCATTAATGACCACTTCCTTTAACTGCCAAGAAAT 1500
DB 1441 GGAGCCGGTAACTGCTCTGCTCTGCTCATTAATGACCACTTCCTTTAACTGCCAAGAAAT 1500
QY 1501 TTTTAAATTAATTAATTTTAAAT 1522
DB 1501 TTTTAAATTAATTAATTTTAAAT 1522

RESULT 3
AAAS2459
ID AAAS2459 standard; cDNA; 1522 BP.
XX
AC AAAS2459;
XX
DT 25-SEP-2000 (first entry)
XX
DE Human MN cDNA.
XX
KW MN protein; tumour associated cell adhesion molecule; oncoprotein;
KW proteoglycan domain; PG domain; carbonic anhydrase; CA domain;
KW abnormal expression; neoplastic disease; cancer; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 13..1392
FT CDS /*tag=a
FT /product="Human MN protein"
FT
XX WO200024913-A2.
XX
PD 04-MAY-2000.
XX
PF 22-OCT-1998; 99WO-US024879.
XX
PR 23-OCT-1998; 98US-00177776.
XX
PR 23-OCT-1998; 98US-00178115.
XX
PA (FARB ) BAYER CORP.
PA (VIRO-) INST VIROLOGY.
XX
PI Zavada J, Pastorekova S, Pastorek J;
XX
DR WPI: 2000-350752/30.
XX
DR P-PSDB; AAB03005.
XX
PT A molecule which specifically binds to a site on MN protein (oncoprotein)
PT and prevents adhesion of vertebrate cells to the protein, useful for
PT treating preneoplastic or neoplastic diseases such as cancer.
XX
PS Example 1; Fig 1A-C; 154pp; English.
XX
CC The invention relates to the inhibition of cell adhesion mediated by the
CC MN oncoprotein (also known as the MN/CA IX isoenzyme or the MN/G250
CC protein). The MN protein is a tumour-associated adhesion molecule which
CC comprises a proteoglycan-like (PG) domain (AA03017) which contains the
CC protein's binding site, and a carbonic anhydrase (CA) domain (AA03018).
CC Abnormal expression of the MN protein is associated with tumorigenicity.
CC The invention encompasses molecules (e.g., proteins and peptides) which
CC which specifically bind to a site on the MN protein, thereby preventing
CC adhesion of vertebrate cells to the protein in a cell adhesion assay. It
CC also encompasses MN proteins or MN protein fragments which can be added
CC to the extracellular environment to prevent the adhesion of vertebrate
CC cells to each other. The invention also relates to the identification of
CC the binding site of the MN protein and to a method of identifying a site
CC on an MN protein to which cells adhere, comprising testing a series of
CC overlapping peptides from the protein in a cell adhesion assay. The
```

```
CC invention encompasses a vector comprising an expression control sequence
CC operatively linked to a nucleic acid encoding the variable domains of a
CC MN-specific antibody, where the domains are separated by a flexible
CC linker peptide (AA03035) and the vector inhibits the growth of a
CC vertebrate preneoplastic or neoplastic cell that abnormally expresses MN
CC protein. The invention also encompasses a vector comprising a nucleic
CC acid encoding a cytotoxic protein or peptide operatively linked to the
CC MN gene promoter, which inhibits the growth of a vertebrate preneoplastic or
CC neoplastic cell. Also claimed is a repressor complex that binds to the MN
CC gene promoter (AAAS2473). MN proteins and peptides, MN-binding proteins
CC and peptides, and expression vectors encoding such proteins and peptides
CC are useful for treating patients with preneoplastic or neoplastic disease
CC (e.g., cancers) associated with or characterized by abnormal MN
CC expression. The present sequence represents cDNA encoding the human MN
CC protein
XX
SQ Sequence 1522 BP; 297 A; 458 C; 453 G; 314 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 1522; DB 3; Length 1522;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACACTAGAGCCGATGAGCTCCCTGTCGCCACGCCCTGAGCTCCCTGTGATCCCGGCC 60
DB 1 ACACTAGAGCCGATGAGCTCCCTGTCGCCACGCCCTGAGCTCCCTGTGATCCCGGCC 60
QY 61 CTTGCTCCAGGCTCACTGTCATGCTGCTGTCATGCTGCTTCTGATGCTTCCAT 120
DB 61 CTTGCTCCAGGCTCACTGTCATGCTGCTGTCATGCTGCTTCTGATGCTTCCAT 120
QY 121 CCCCAGAGGTTGCCCCGGATGCGAGAGATCCCTTGGAGAGAGCTCTTGGGAA 180
DB 121 CCCCAGAGGTTGCCCCGGATGCGAGAGATCCCTTGGAGAGAGCTCTTGGGAA 180
QY 121 CCCCAGAGGTTGCCCCGGATGCGAGAGATCCCTTGGAGAGAGCTCTTGGGAA 180
DB 121 CCCCAGAGGTTGCCCCGGATGCGAGAGATCCCTTGGAGAGAGCTCTTGGGAA 180
QY 181 GATGATCCCACTGGGCGAGAGAGATCTGCCAGTGAAGAGATTCAACCAAGAGAGAT 240
DB 181 GATGATCCCACTGGGCGAGAGAGATCTGCCAGTGAAGAGATTCAACCAAGAGAGAT 240
QY 241 CCACCCGAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 300
DB 241 CCACCCGAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 300
QY 301 GAAATTAAAGCTTAATCAAGAGAGAGGCTCTTGAAGTAGAGATCTACTACTGTT 360
DB 301 GAAATTAAAGCTTAATCAAGAGAGAGGCTCTTGAAGTAGAGATCTACTACTGTT 360
QY 361 GAGGCTCTGAGATCTCTCAAGAACCCGAAATTAATGCCACAGGAGCAAGAGAGAT 420
DB 361 GAGGCTCTGAGATCTCTCAAGAACCCGAAATTAATGCCACAGGAGCAAGAGAGAT 420
QY 421 GACCAAGATATTGAGGCTATGAGGAGACCCGCTGAGCCCGGATGTCGCCAGCTGC 480
DB 421 GACCAAGATATTGAGGCTATGAGGAGACCCGCTGAGCCCGGATGTCGCCAGCTGC 480
QY 481 GACCAAGATATTGAGGCTATGAGGAGACCCGCTGAGCCCGGATGTCGCCAGCTGC 480
DB 481 GACCAAGATATTGAGGCTATGAGGAGACCCGCTGAGCCCGGATGTCGCCAGCTGC 480
QY 481 GCGGCGGCTTCCAGTCCCGGATGATCCCGCCCAAGCTGCGGCTTCTGCGCGGC 540
DB 481 GCGGCGGCTTCCAGTCCCGGATGATCCCGCCCAAGCTGCGGCTTCTGCGCGGC 540
QY 541 CTGGGCCCCCTGAGATCTCTGAGCTTCCAGCTCCGCGCTCCAGAACTGCGCTGCGC 600
DB 541 CTGGGCCCCCTGAGATCTCTGAGCTTCCAGCTCCGCGCTCCAGAACTGCGCTGCGC 600
QY 601 AACATATGCGACAGTGTGACATGACCTGCTCTTGGGCTGAGATGCTTGGGATCC 660
DB 601 AACATATGCGACAGTGTGACATGACCTGCTCTTGGGCTGAGATGCTTGGGATCC 660
QY 661 GAGCGGAGATACCGGGCTGAGAGCTGATGCACTGAGGAGGCTGCAAGATGTCGAGC 720
DB 661 GAGCGGAGATACCGGGCTGAGAGCTGATGCACTGAGGAGGCTGCAAGATGTCGAGC 720
QY 721 TCGAGAGACACTGTGAGAGGACCGCTTCTGCTGCGAGATCCAGGTGTTTCACTGAC 780
DB 721 TCGAGAGACACTGTGAGAGGACCGCTTCTGCTGCGAGATCCAGGTGTTTCACTGAC 780
```



```

QY 781 ACCGCTTTCAGAGTTGACAGAGGCTTGGGGCCCGGGAGGCTCGCCCTGTTGGCC 840
DB 781 ACCGCTTTCAGAGTTGACAGAGGCTTGGGGCCCGGGAGGCTCGCCCTGTTGGCC 840
QY 841 GCGTTTCTGAGAGAGGCGCCGAGAGAAACAGTGTCTATGACAGTTGCTGTCTGCTTG 900
DB 841 GCGTTTCTGAGAGAGGCGCCGAGAGAAACAGTGTCTATGACAGTTGCTGTCTGCTTG 900
QY 901 GAAGAAATGCTGAGAGAGAGCTCAGAGACTCAGGTCCAGAGACTGACATATCTGACCTC 960
DB 901 GAAGAAATGCTGAGAGAGAGCTCAGAGACTCAGGTCCAGAGACTGACATATCTGACCTC 960
QY 961 CTGCGCTCTGACTGAGCGGCTACTTCCAAATATGAGGGGTCTGACTACACCGCCCTGT 1020
DB 961 CTGCGCTCTGACTGAGCGGCTACTTCCAAATATGAGGGGTCTGACTACACCGCCCTGT 1020
QY 1021 GCGCAGGCTGATCTGAGACTGTGTTAACAGAGAGTGAATGCTGAGTGTCTAGAGACTC 1080
DB 1021 GCGCAGGCTGATCTGAGACTGTGTTAACAGAGAGTGAATGCTGAGTGTCTAGAGACTC 1080
QY 1081 CACACCTCTCTGACACCTGTGGGACCTGTGACTCTCGGCTACAGACTGAATTTCCGA 1140
DB 1081 CACACCTCTCTGACACCTGTGGGACCTGTGACTCTCGGCTACAGACTGAATTTCCGA 1140
QY 1141 GCGAGCGAGCCTTTGAAATGAGGCGAGTGAATGAGGCTCTCTCTGCTGAGTGGACAGC 1200
DB 1141 GCGAGCGAGCCTTTGAAATGAGGCGAGTGAATGAGGCTCTCTCTGCTGAGTGGACAGC 1200
QY 1201 AGTCTTGCGGCTGCTGAGCGAGTGAATGCTGAGTGTCTGCTGCTGCTGAGTGAATCTTA 1260
DB 1201 AGTCTTGCGGCTGCTGAGCGAGTGAATGCTGAGTGTCTGCTGCTGCTGAGTGAATCTTA 1260
QY 1261 GCGCTGCTTTTGGCTCTCTTTTGTCTGTCACAGCGTGGCGTCTCTGCTGAGTGAATGAGA 1320
DB 1261 GCGCTGCTTTTGGCTCTCTTTTGTCTGTCACAGCGTGGCGTCTCTGCTGAGTGAATGAGA 1320
QY 1321 AGGACAGACAGAGAGGAGAACAAAGAGGCTGTGAGTCAACGCGCAGAGAGAGTGGCCGAG 1380
DB 1321 AGGACAGACAGAGAGGAGAACAAAGAGGCTGTGAGTCAACGCGCAGAGAGAGTGGCCGAG 1380
QY 1381 ACTGAGAGCTAGAGGCTGAGTCTTGAGAGATGTGAGAGAGCAGAGGAGCATCTGAGGG 1440
DB 1381 ACTGAGAGCTAGAGGCTGAGTCTTGAGAGATGTGAGAGAGCAGAGGAGCATCTGAGGG 1440
QY 1441 GAGAGCGGCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
DB 1441 GAGAGCGGCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
QY 1501 TTTTAAATTAATATTTATTAAT 1522
DB 1501 TTTTAAATTAATATTTATTAAT 1522

```

```

RESULT 4
ABL64779
ID ABL64779 standard; DNA; 1552 BP.
XX
AC ABL64779;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3116.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN MO200194629-A2.
XX

```

```

PD 13-DEC-2001.
XX
PR 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-023133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237608P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 3116; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in (1)
XX expression of at least one gene (1) of a signature gene set, where (1)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (1) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result
XX of M1, and the data is sufficient to convey the chemical structure and/or
XX properties of the agent. M1 can be used in the treatment of cancer such
XX as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,

```


PR 22-SEP-2000; 2000US-0234503P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237585P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 516; 44bp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
CC tumour
XX
SQ Sequence 1552 BP; 302 A; 471 C; 461 G; 318 T; 0 U; 0 Other;
Query Match 100.0%; Score 1522; DB 6; Length 1552;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGTCAAGCCGATGAGTCCCTGCTGTCGCCAGCCCTGAGTCCCTCTGTGATCCCGGCC 60
Db 31 AAGTCAAGCCGATGAGTCCCTGCTGTCGCCAGCCCTGAGTCCCTCTGTGATCCCGGCC 90
QY 61 CTTGCTCCAGGCTCCTCATGTGCACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 91 CTTGCTCCAGGCTCCTCATGTGCACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150
QY 121 CCCAGAGGTTGCCCGGATGACAGAGAGATTCCTCCCTTGGAGAGAGCTCTTGGGGA 180
Db 151 CCCAGAGGTTGCCCGGATGACAGAGAGATTCCTCCCTTGGAGAGAGCTCTTGGGGA 210
QY 181 GATGACCACTGGGCGAGAGAGATCTGCTCCAGTAAGAGATTCACCCAGAGAGAGAT 240
Db 211 GATGACCACTGGGCGAGAGAGATCTGCTCCAGTAAGAGATTCACCCAGAGAGAGAT 270
QY 241 CCACCCGAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGAGATCTACCT 300
Db 271 CCACCCGAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 330
QY 301 GAAGTTAAGCTTAATATGAAAGAGAGGCTCCCTGAGTTAAGAGATCTACCTACTGTT 360
Db 331 GAAGTTAAGCTTAATATGAAAGAGAGGCTCCCTGAGTTAAGAGATCTACCTACTGTT 390
QY 361 GAGGCTCCTGAGATCTCTCAAGAAAGAGAGATTAATGAGAGAGAGAGAGAGAT 420
Db 391 GAGGCTCCTGAGATCTCTCAAGAAAGAGAGATTAATGAGAGAGAGAGAGAGAT 450
QY 421 GAACAGAGTCAATTGGGCTATGAGAGGAGACCCGCTGAGCTCCGAGTCTCCAGCTG 480
Db 451 GAACAGAGTCAATTGGGCTATGAGAGGAGACCCGCTGAGCTCCGAGTCTCCAGCTG 510
QY 481 GCGGCGGCTTCCGAGTCCCGGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db 511 GCGGCGGCTTCCGAGTCCCGGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 570
QY 541 CTGGCGCCCTGGAATCTGAGGCTTCCAGCTCCCGGCTCCAGAACTGCGCTGCGC 600
Db 571 CTGGCGCCCTGGAATCTGAGGCTTCCAGCTCCCGGCTCCAGAACTGCGCTGCGC 630
QY 601 AACCAATGCGCACTGTGCAACTGACCTTCCCTGAGGCTAGAGATGAGTGGGATGCC 660
Db 631 AACCAATGCGCACTGTGCAACTGACCTTCCCTGAGGCTAGAGATGAGTGGGATGCC 690
QY 661 GGGCGGAGATCCCGGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATG 720
Db 691 GGGCGGAGATCCCGGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATG 750
QY 721 TCGAGGACACTGTGGAAGGCAACCGCTTCCCTGCGAATCCACGCTGCTACCTGAC 780
Db 751 TCGAGGACACTGTGGAAGGCAACCGCTTCCCTGCGAATCCACGCTGCTACCTGAC 810
QY 781 ACCGCTTGGCAGAGTTGACGAGGCTTGGGCGCCCGAGAGGCTGCGCTGTTGGCC 840
Db 811 ACCGCTTGGCAGAGTTGACGAGGCTTGGGCGCCCGAGAGGCTGCGCTGTTGGCC 870
QY 841 GCTTTTCTGAGAGGCGCCGGAAGAAACAGTCCCTATGACAGATGCTGCTGCTG 900
Db 871 GCTTTTCTGAGAGGCGCCGGAAGAAACAGTCCCTATGACAGATGCTGCTGCTG 930
QY 901 GAAGAAATGCTGAGAGAGGCTGAGAGCTCAGATCCAGAGCTGACATATGCACTC 960
Db 931 GAAGAAATGCTGAGAGAGGCTGAGAGCTCAGATCCAGAGCTGACATATGCACTC 990
QY 961 CTGGCTTGAATCTGAGCGCTGACTTCCATATGAGAGGCTGCTGACTAACCCTGCT 1020
Db 991 CTGGCTTGAATCTGAGCGCTGACTTCCATATGAGAGGCTGCTGACTAACCCTGCT 1050
QY 1021 GCCCAGGATGATCTGAGCTGTGTTTAAACAGACAGTGAAGTGAAGTGAAGTGAAG 1080
Db 1051 GCCCAGGATGATCTGAGCTGTGTTTAAACAGACAGTGAAGTGAAGTGAAGTGAAG 1110
QY 1081 CACACCTCTGTAACACCTGAGGAGCTGTGAGCTCTGAGCTACAGTGAACCTTCCGA 1140

QY 181 GATGACCACTGGGAGAGAGATCTGCCAGTGAAGATTCACCCAGAGAGAGAT 240
DB 211 GATGACCACTGGGAGAGAGATCTGCCAGTGAAGATTCACCCAGAGAGAGAT 270
QY 241 CCACCCGAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 300
DB 271 CCACCCGAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGATCTACCT 330
QY 301 GAAGTTAGCTTAATCAAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAT 360
DB 331 GAAGTTAGCTTAATCAAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAT 390
QY 361 GAGGCTCTGAGAGATCTCAAGAACCCCAAGATTAATGCCCACAGGGAACAAAGAGGAGAT 420
DB 391 GAGGCTCTGAGAGATCTCAAGAACCCCAAGATTAATGCCCACAGGGAACAAAGAGGAGAT 450
QY 421 GACCAAGTCAATGGGGCTATGAGAGGACCCCGCTGAGAGAGATCTACCTGAGAGAT 480
DB 451 GACCAAGTCAATGGGGCTATGAGAGGACCCCGCTGAGAGAGATCTACCTGAGAGAT 510
QY 481 GCGGAGCGCTTCAGATCCCGAGTATCCGAGATCCCGAGATCCCGAGATCCCGAGAT 540
DB 511 GCGGAGCGCTTCAGATCCCGAGTATCCGAGATCCCGAGATCCCGAGATCCCGAGAT 570
QY 541 CTGGGCGCTTCAGATCCCGAGTATCCGAGATCCCGAGATCCCGAGATCCCGAGAT 600
DB 571 CTGGGCGCTTCAGATCCCGAGTATCCGAGATCCCGAGATCCCGAGATCCCGAGAT 630
QY 601 AACATGAGCAGAGTGTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 660
DB 631 AACATGAGCAGAGTGTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 690
QY 661 GGGCGGAGTACCGGCTCTGAGATCTGAGATCTGAGATCTGAGATCTGAGATCTGAGAT 720
DB 691 GGGCGGAGTACCGGCTCTGAGATCTGAGATCTGAGATCTGAGATCTGAGATCTGAGAT 750
QY 721 TCGAGACACACTGTGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 780
DB 751 TCGAGACACACTGTGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 810
QY 781 ACCGCTTTTCCAGATGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 840
DB 811 ACCGCTTTTCCAGATGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 870
QY 841 GCGCTTCTGAGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 900
DB 871 GCGCTTCTGAGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 930
QY 901 GAAGAAATGCTGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 960
DB 931 GAAGAAATGCTGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 990
QY 961 CTGGCGCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1020
DB 991 CTGGCGCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1050
QY 1021 GCGCAGAGTGTGATCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1080
DB 1051 GCGCAGAGTGTGATCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1110
QY 1081 GACACCTCTCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1140
DB 1111 GACACCTCTCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1170
QY 1141 GCGCAGAGTGTGATCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1200
DB 1171 GCGCAGAGTGTGATCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1230
QY 1201 AGTCTCTGGGCTGCTGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1260
DB 1231 AGTCTCTGGGCTGCTGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1290

QY 1261 GCCCTGTTTGGGCTCTTTTGTCTGACAGCCGTCGCTTCTTGTGAGATGAGA 1320
DB 1291 GCCCTGTTTGGGCTCTTTTGTCTGACAGCCGTCGCTTCTTGTGAGATGAGA 1350
QY 1321 AGGCAGCAGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1380
DB 1351 AGGCAGCAGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1410
QY 1381 ACTGAGCTTGAAGCTGATCTTGAAGATGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1440
DB 1411 ACTGAGCTTGAAGCTGATCTTGAAGATGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1470
QY 1441 GAGCCGCTTGAAGCTGATCTTGAAGATGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1500
DB 1471 GAGCCGCTTGAAGCTGATCTTGAAGATGAGAGAGAGATCCCTGAGATCCCTGAGATCCCTGAGAT 1530
QY 1501 TTTTAAATTAATTTAAT 1552
DB 1531 TTTTAAATTAATTTAAT 1552

RESULT 7
ABL68346
ID ABL68346 standard; DNA; 1552 BP.
XX
AC ABL68346;
XX
DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:683.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
gene; ds.
XX
OS Homo sapiens.
XX
FN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0231137P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 25-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 26-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.

PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237284P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 02-OCT-2000; 2000US-0237345P.
PR 02-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237608P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.

(AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 6683; 44pp; English.

XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (1) of a signature gene set, where (1)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL1664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (1) has cytosolic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result
XX of M1, and the data is sufficient to convey the chemical structure and/or
XX properties of the agent. M1 can be used in the treatment of cancer such
XX as colon, breast, stomach, lung, thyroid, esophageal, ovarian, kidney,
XX prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
XX cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
XX tumour

XX Sequence 1552 BP; 302 A; 471 C; 461 G; 318 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 1552; DB 6; Length 1552;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 1552; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGTCAGCCGATGCTCCCTGTGCCCAAGCCCTGCTCCTCTGTTGATCCGAGCC 60
DB 31 ACAGTCAGCCGATGCTCCCTGTGCCCAAGCCCTGCTCCTCTGTTGATCCGAGCC 90
QY 61 CCTGCTCCAGGCTCTCACTGTGTGCACTGTCTGTGCTGCTGCTTGTGATGCTTCAT 120
DB 91 CCTGCTCCAGGCTCTCACTGTGTGCACTGTCTGTGCTGCTGCTTGTGATGCTTCAT 150
QY 121 CCCGAGAGTTGCTCCCGGATGTCAGAGAGATTCCTCCCTTGAGAGAGGCTCTTCTGAGAA 180
DB 151 CCCGAGAGTTGCTCCCGGATGTCAGAGAGATTCCTCCCTTGAGAGAGGCTCTTCTGAGAA 210
QY 181 GATGACCACTGAGGAGAGAGATTCGCCAGTGAAGAGATTACCCAGAGAGAGAT 240
DB 211 GATGACCACTGAGGAGAGAGATTCGCCAGTGAAGAGATTACCCAGAGAGAGAT 270
QY 241 CCACCCGAGAGAGAGATTCACCTGAGAGAGATTCACCTGAGAGAGAGATTCACCT 300
DB 271 CCACCCGAGAGAGAGATTCACCTGAGAGAGATTCACCTGAGAGAGAGATTCACCT 330
QY 301 GAAGTTAAGCTTAATCAGAGAGAGAGGCTCCTGAGAGTTAGAGATTCACCTGTT 360

DB 331 GAAGTTAAGCTTAATCAGAGAGAGAGGCTCCTGAGAGTTAGAGATTCACCTGTT 390
QY 361 GAGGCTCTGAGAGATTCCTAGAGACCCCAAGTAATATGCCACAGGAGCAAGAGAGAT 420
DB 391 GAGGCTCTGAGAGATTCCTAGAGACCCCAAGTAATATGCCACAGGAGCAAGAGAGAT 450
QY 421 GACCAAGATCATTGCGCTATGAGAGCGACCCGCTTGGCCCGGAGTGTCCCAAGCTTC 480
DB 451 GACCAAGATCATTGCGCTATGAGAGCGACCCGCTTGGCCCGGAGTGTCCCAAGCTTC 510
QY 481 GCGGCGCGCTTCAGTCCCGGAGTATCCGCGCCCAAGTGGCGGCTTCTGCGCGGAGC 540
DB 511 GCGGCGCGCTTCAGTCCCGGAGTATCCGCGCCCAAGTGGCGGCTTCTGCGCGGAGC 570
QY 541 CTGCGCCCTCTGGAATCCTGAGGCTTCCAGCTCCGCGCTTCCAGAACTGCGCTGCGC 600
DB 571 CTGCGCCCTCTGGAATCCTGAGGCTTCCAGCTCCGCGCTTCCAGAACTGCGCTGCGC 630
QY 601 AACAAATGCGACAGTGTGCACTGACCCCTGCTGCTGAGAGATGCTGAGGCTTCC 660
DB 631 AACAAATGCGACAGTGTGCACTGACCCCTGCTGCTGAGAGATGCTGAGGCTTCC 690
QY 661 GGGGAGGAGTACCGGCTCTGACGCTGATCTGCACTGAGGAGGCTGCAAGTCTCGGAGC 720
DB 691 GGGGAGGAGTACCGGCTCTGACGCTGATCTGCACTGAGGAGGCTGCAAGTCTCGGAGC 750
QY 721 TCGAGAGACATCTGTGAAGGACCGCTTCCGCGGAGATCCAGTGTGCTACCTGAC 780
DB 751 TCGAGAGACATCTGTGAAGGACCGCTTCCGCGGAGATCCAGTGTGCTACCTGAC 810
QY 781 ACCGCTTTTCCAGAGTTGACAGAGGCTTGGAGGCGCGGAGGCTGAGGCTGTTGAGC 840
DB 811 ACCGCTTTTCCAGAGTTGACAGAGGCTTGGAGGCGCGGAGGCTGAGGCTGTTGAGC 870
QY 841 GCTTTTGTGAGAGAGGCGCGGAGAGAAACAGTGCCTATGACAGTGTGCTGAGCTTG 900
DB 871 GCTTTTGTGAGAGAGGCGCGGAGAGAAACAGTGCCTATGACAGTGTGCTGAGCTTG 930
QY 901 GAAGAATTCCTGAGAGAGGCTGAGAGTCCAGAGTCCAGAGCTGAGATCTGACATC 960
DB 931 GAAGAATTCCTGAGAGAGGCTGAGAGTCCAGAGTCCAGAGCTGAGATCTGACATC 990
QY 961 CTGCGCTCTGATTCAGCGGCTATCTTCATATGAGGCTCTGATACAGCGGCTGT 1020
DB 991 CTGCGCTCTGATTCAGCGGCTATCTTCATATGAGGCTCTGATACAGCGGCTGT 1050
QY 1021 GCCGAGGCTGATCTGAGATCTGTGTTAACAGACAGATGATGCTGATGAGAGCTC 1080
DB 1051 GCCGAGGCTGATCTGAGATCTGTGTTAACAGACAGATGATGCTGATGAGAGCTC 1110
QY 1081 CACACCTCTCTGACACCTGTGTGGGACCTGTGACTCTCGGCTACAGTGAATCTCCGA 1140
DB 1111 CACACCTCTCTGACACCTGTGTGGGACCTGTGACTCTCGGCTACAGTGAATCTCCGA 1170
QY 1141 GCGAGCGAGGCTTGAATGAGGAGTGAAGGCTCTCTTCCCTGAGAGAGAGCT 1200
DB 1171 GCGAGCGAGGCTTGAATGAGGAGTGAAGGCTCTCTTCCCTGAGAGAGAGCT 1230
QY 1201 AGTCTTGGGCTGAGGAGTGAAGTGAAGGCTCTCTTCCCTGAGAGAGAGCT 1260
DB 1231 AGTCTTGGGCTGAGGAGTGAAGTGAAGGCTCTCTTCCCTGAGAGAGAGCT 1290
QY 1261 GCGCTGATTTTGGCTCTCTTTTGTGCTGACACAGGCTGCGCTCTCTTGTGAGATGAGA 1320
DB 1291 GCGCTGATTTTGGCTCTCTTTTGTGCTGACACAGGCTGCGCTCTCTTGTGAGATGAGA 1350
QY 1321 AGGACGACACAGAGGAGACCAAAAGGAGTGTAGCTACCCGCGACAGAGTACCGGAG 1380
DB 1351 AGGACGACACAGAGGAGACCAAAAGGAGTGTAGCTACCCGCGACAGAGTACCGGAG 1410
QY 1381 ACTGAGGCTTGAAGGCTGATCTTGAAGAACTGAGAGGCTGAGAGGCTGAGAGG 1440
DB 1411 ACTGAGGCTTGAAGGCTGATCTTGAAGAACTGAGAGGCTGAGAGGCTGAGAGG 1470

QY 1441 GGAGCCGTAACCTGCTGCTGCTCATATGACACCTTCTTTAACTCCAGAAATT 1500
 Db 1471 GGAGCCGTAACCTGCTGCTGCTCATATGACACCTTCTTTAACTCCAGAAATT 1530
 QY 1501 TTTTAAATTAATATTTTAAAT 1522
 Db 1531 TTTTAAATTAATATTTTAAAT 1552
 RESULT 8
 ACC72730
 ID ACC72730 standard; cDNA; 1552 BP.
 XX ACC72730;
 AC ACC72730;
 XX 09-JUL-2003 (first entry)
 DT 09-JUL-2003 (first entry)
 XX Human cancer related protein encoding cDNA SEQ ID NO:69.
 DE Human cancer related protein encoding cDNA SEQ ID NO:69.
 XX Human; cancer; diagnosis; screening; modulator; leukemia; ischaemia;
 KW heart disease; atherosclerosis; endometrios; gene; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 PN NC02003025138-A2.
 XX NC02003025138-A2.
 PD 27-MAR-2003.
 XX 17-SEP-2002; 2002WC-US029560.
 PF 17-SEP-2001; 2001US-0323469P.
 PR 20-SEP-2001; 2001US-0323469P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 08-FEB-2002; 2002US-0355145P.
 PR 08-FEB-2002; 2002US-0355257P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (E0SB) - EOS BIOTECHNOLOGY INC.
 PA Afar D, Aziz N, Gish KC, Hevez PA, Mack DH, Wilson KE;
 PI Zlotnick A;
 PT WPI; 2003-354600/33.
 DR P-PSDB; ABR58595.
 XX New genes that are up-regulated or down-regulated in cancers, useful as
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 PT therapeutic targets for screening drugs for treating these diseases.
 PS Claim 8; Page 662; 767P; English.
 CC The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
 CC acute lymphocytic leukemia). ACC72641 to ACC72880 represent cancer
 CC related gene nucleotide sequences which encode the proteins given in
 CC ABR58521 to ABR58709. Also described: (1) determining the presence or
 CC absence of a pathological cell in a patient; (2) an expression vector
 CC comprising a nucleic acid molecule described above; (3) a host cell
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
 CC of (4); (6) specifically targeting a compound to a pathological cell in a
 CC patient by administering to the patient the nucleic acid is useful as
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or
 CC therapeutic targets. In particular, the nucleic acid is useful for
 CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
 CC pancreas, prostate, skin and uterus), wounds, ischemia, heart diseases,
 CC atherosclerosis and endometrios. The nucleic acid is also useful in
 CC drug screening, particularly for identifying agents for treating these
 CC pathologies
 XX

SQ Sequence 1552 BP; 302 A; 471 C; 461 G; 318 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1522; DB 7; Length 1552;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACAGTCAGCCGATGAGCTCCCTGTCGCGCCAGCCCTGCGCTCTGTGATCCCGGCC 60
 Db 31 ACAGTCAGCCGATGAGCTCCCTGTCGCGCCAGCCCTGCGCTCTGTGATCCCGGCC 90
 QY 61 CCGTCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 Db 91 CCGTCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150
 QY 121 CCCGAGAGTTGGCCCGGAGTGCAGAGAGATTTCCCTGAGAGAGAGCTTTCTGAGGAA 180
 Db 151 CCCGAGAGTTGGCCCGGAGTGCAGAGAGATTTCCCTGAGAGAGAGCTTTCTGAGGAA 210
 QY 181 GATGACCCGCTGAGGCGAGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGAT 240
 Db 211 GATGACCCGCTGAGGCGAGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGAT 270
 QY 241 CCACCCGAGAGAGAGATCTGCTGAGAGAGAGATCTGCTGAGAGAGAGATCTGCTGCT 300
 Db 271 CCACCCGAGAGAGAGATCTGCTGAGAGAGAGATCTGCTGAGAGAGAGATCTGCTGCT 330
 QY 301 GAAGTTAAGCTTAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCT 360
 Db 331 GAAGTTAAGCTTAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCT 390
 QY 361 GAGGCTCTGAGAGATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 420
 Db 391 GAGGCTCTGAGAGATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 450
 QY 421 GACGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCTG 480
 Db 451 GACGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCTG 510
 QY 481 GCGGCGCGCTTCACTCCCGGTGATATCCGCGCCAGAGCTGCGCGCTTCTGCGCGGC 540
 Db 511 GCGGCGCGCTTCACTCCCGGTGATATCCGCGCCAGAGCTGCGCGCTTCTGCGCGGC 570
 QY 541 CTGCGCGCGCTTCACTCCCGGTGATATCCGCGCCAGAGCTGCGCGCTTCTGCGCGGC 600
 Db 571 CTGCGCGCGCTTCACTCCCGGTGATATCCGCGCCAGAGCTGCGCGCTTCTGCGCGGC 630
 QY 601 AACAAATGGCAAGTGTGCAACTGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 631 AACAAATGGCAAGTGTGCAACTGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 690
 QY 661 GGGCGGAGTACCGGGCTCTGAGAGTGTGCACTGAGAGAGAGAGAGAGAGAGAGAG 720
 Db 691 GGGCGGAGTACCGGGCTCTGAGAGTGTGCACTGAGAGAGAGAGAGAGAGAGAGAG 750
 QY 721 TCGAGACACTGTGAAAGGCCACCGTTTCCCTGCGAGATCCACGTGTTCACCTAGC 780
 Db 751 TCGAGACACTGTGAAAGGCCACCGTTTCCCTGCGAGATCCACGTGTTCACCTAGC 810
 QY 781 ACCGCTTTCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 Db 811 ACCGCTTTCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 870
 QY 841 GCGTTTCTGAG 900
 Db 871 GCGTTTCTGAG 930
 QY 901 GAAGAAATGCTGAG 960
 Db 931 GAAGAAATGCTGAG 990
 QY 961 CTGCGCTTCACTTACCGCGCTTCACTTCAATATGAGAGAGAGAGAGAGAGAGAGAG 1020
 Db 991 CTGCGCTTCACTTACCGCGCTTCACTTCAATATGAGAGAGAGAGAGAGAGAGAGAG 1050

QY 1021 GCCCAGGATGTCATCTGACTGTGTTTAAACCAAGATGATGCTGAATGCTTAAGCAGCTC 1080
 DB 1051 GCCCAGGATGTCATCTGACTGTGTTTAAACCAAGATGATGCTGAATGCTTAAGCAGCTC 1110
 QY 1081 CACACCCCTCTGACACACCTGTGGGAGCTGTGATCTGGGCTACAGCTGAATCCGA 1140
 DB 1111 CACACCCCTCTGACACACCTGTGGGAGCTGTGATCTGGGCTACAGCTGAATCCGA 1170
 QY 1141 GCGACGACCCCTTTGAATGGGAGATGATGAGCCCTTCCCTGCTGAGTGAAGCAGC 1200
 DB 1171 GCGACGACCCCTTTGAATGGGAGATGATGAGCCCTTCCCTGCTGAGTGAAGCAGC 1230
 QY 1201 ACTGCTCGGGCTGTGACGACGCTGAGTGAATCTGAGCTGAGTGTGATGATGATGAT 1260
 DB 1231 ACTGCTCGGGCTGTGACGACGCTGAGTGAATCTGAGCTGAGTGTGATGATGATGAT 1290
 QY 1261 GCGCTGATTTTGGGCTCTTTTGGCTGTGACGACGCTGAGTGTGATGATGATGATGAT 1320
 DB 1291 GCGCTGATTTTGGGCTCTTTTGGCTGTGACGACGCTGAGTGTGATGATGATGATGAT 1350
 QY 1321 AGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1380
 DB 1351 AGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1410
 QY 1381 ACTGAGCCTTGAAGGCTGATCTTGAAGATGTGAGAGCCAGCAGAGGATCTGAGAG 1440
 DB 1411 ACTGAGCCTTGAAGGCTGATCTTGAAGATGTGAGAGCCAGCAGAGGATCTGAGAG 1470
 QY 1441 GAGCCGGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 DB 1471 GAGCCGGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1530
 QY 1501 TTTTAAATTAATTAATTAATTAAT 1522
 DB 1531 TTTTAAATTAATTAATTAATTAAT 1552

RESULT 9
 AB277284
 ID AB277284 standard; cDNA, 1552 BP.
 XX
 AC AB277284;
 DT 28-MAY-2003 (first entry)
 XX
 DE Nucleotide sequence of human CA9 polypeptide.
 XX
 KW Human; angiogenesis; wound healing; retinopathy; ischemia; inflammation;
 KW microvasculopathy; bone healing; skin inflammation; HOG3; HOG3; HOG18;
 KW follicular development; CA9; HXB; IGFBP5; HEPAR; STC1; miG-6; SSR4;
 KW cancer; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 43..1422
 FT /tag=a
 FT /product="CA9"
 XX
 PN WO2003010205-A1.
 XX
 PD 06-FEB-2003.
 XX
 PF 26-JUL-2002; 2002WO-US023786.
 XX
 PR 26-JUL-2001; 2001US-030760P.
 PR 24-JUL-2002; 2002US-00201642.
 XX
 PA (UYDU-) UNIV DUKE MEDICAL CENT.
 XX
 PI Rigsins GJ, Lal A;
 XX

DR WPI: 2003-239423/23.
 DR P-ESDB; ABP9744.
 XX
 XX Inhibiting angiogenesis for treating wound healing, retinopathy,
 PT ischemia, inflammation, microvasculopathy, bone healing, skin
 PT inflammation or follicular development by providing to a subject an
 PR antisense polynucleotide.
 XX
 PS Claim 1; Page 47; 66pp; English.
 CC
 CC The present sequence encodes a human CA9 polypeptide. It is used in the
 CC method of the invention. The specification describes a method of modulating
 CC angiogenesis associated with wound healing, retinopathy, ischemia,
 CC inflammation, microvasculopathy, bone healing, skin inflammation or
 CC follicular development. The method comprises providing to a subject HOG3,
 CC HOG8, HOG18, CA9, HXB, IGFBP5, HEPAR, STC1, miG-6 or SSR4. The methods,
 CC antisense polynucleotides, polypeptides and antibodies are useful for
 CC treating wound healing, retinopathy, ischemia, inflammation,
 CC microvasculopathy, bone healing, skin inflammation or follicular
 CC development, or cancer such as breast, colon or lung cancer, or
 CC glioblastoma
 CC
 SQ Sequence 1552 BP; 302 A; 471 C; 461 G; 318 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1522; DB 7; Length 1552;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACAGTCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 DB 31 ACAGTCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 90
 QY 61 CCGTGTCCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 DB 91 CCGTGTCCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 150
 QY 121 CCGCAGAGGTTGCCCGGATGACAGAGATTTCCCTTGGAGAGAGCTTTCTGGAGAA 180
 DB 151 CCGCAGAGGTTGCCCGGATGACAGAGATTTCCCTTGGAGAGAGCTTTCTGGAGAA 210
 QY 181 GATGACCATGCGGCGAGAGAGATCTGCGCAGTGAAGAGATTCACCCAGAGAGAGAT 240
 DB 211 GATGACCATGCGGCGAGAGAGATCTGCGCAGTGAAGAGATTCACCCAGAGAGAGAT 270
 QY 241 CCACCCGAGAGAGAGATCTGCTGAGAGAGATCTGCTGAGAGAGATCTTAACT 300
 DB 271 CCACCCGAGAGAGAGATCTGCTGAGAGAGATCTGCTGAGAGAGATCTTAACT 330
 QY 301 GAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGAAGTTAAGAGATCTTAACT 360
 DB 331 GAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGAAGTTAAGAGATCTTAACT 390
 QY 361 GAGGCTCTGAGATCTCCTCAAGAACCCAGATATATCCACAGAGAGCAAGAGAGGAT 420
 DB 391 GAGGCTCTGAGATCTCCTCAAGAACCCAGATATATCCACAGAGAGCAAGAGAGGAT 450
 QY 421 GACCAAGATCTGAGGCTATGAGGAGCCGCTGAGCCCGGGTGTCCCAAGCTTC 480
 DB 451 GACCAAGATCTGAGGCTATGAGGAGCCGCTGAGCCCGGGTGTCCCAAGCTTC 510
 QY 481 GCGGCGGCTTCCAGTCCCGGCTGAGATATGCGCCGAGCTGCGCGCTTCTGCGCGGC 540
 DB 511 GCGGCGGCTTCCAGTCCCGGCTGAGATATGCGCCGAGCTGCGCGCTTCTGCGCGGC 570
 QY 541 CTGCGCCCTCTGAACTCTGAGGCTTCAGCTCCCGCGCTCCCAAGATCTGCGCGGC 600
 DB 571 CTGCGCCCTCTGAACTCTGAGGCTTCAGCTCCCGCGCTCCCAAGATCTGCGCGGC 630
 QY 601 AACATGCGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
 DB 631 AACATGCGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 690
 QY 661 GGGCGGAGTACCGGGCTCTGAGAGCTGATCTGACTGGGGGGCTGAGAGTCTGCGGGC 720

QY	24	CCACC	CGAGAGAGAGATCTTCACTGGAGAGAGATCTTACTGAGAGAGATCTTACT	300
Db	271	CCACC	CGAGAGAGAGATCTTCACTGGAGAGAGATCTTACTGAGAGAGATCTTACT	330
QY	301	GAA	TTAAECTTAAATCAAGAAAGAGGCTCCTGAAATTAGAGATTTACTCTGT	360
Db	331	GAA	TTAACTTAAATCAAGAAAGAGGCTCCTGAAATTAGAGATTTACTCTGT	390
QY	361	GAG	GCCTCTGGAGATCTTCAAGAAACCCAGATTAATCCACAGGGACAAAGAGGGAT	420
Db	391	GAG	GCCTCTGGAGATCTTCAAGAAACCCAGATTAATCCACAGGGACAAAGAGGGAT	450
QY	421	GAC	CAGATCATTTGGCGCTATGAGGGGACCCGGCTTGACCCCGGATGTCCCAACCTGC	480
Db	451	GAC	CAGATCATTTGGCGCTATGAGGGGACCCGGCTTGACCCCGGATGTCCCAACCTGC	510
QY	481	GCG	GGCGGCTTCAGTCCCGGATGATATCCGCGCCAGCTCGCCGCTTCTGCGCGGAC	540
Db	511	GCG	GGCGGCTTCAGTCCCGGATGATATCCGCGCCAGCTCGCCGCTTCTGCGCGGAC	570
QY	541	CTG	CGCCCCCTGGAATCTCTGGGCTTCCAGTCTCCCGCTCCAGAACTGCGCTCGC	600
Db	571	CTG	CGCCCCCTGGAATCTCTGGGCTTCCAGTCTCCCGCTCCAGAACTGCGCTCGC	630
QY	601	AAC	AATGAGGACAGATGTGACATAACCTGTGCTCTGGGCTTAGAGATTTGGATCC	660
Db	631	AAC	AATGAGGACAGATGTGACATAACCTGTGCTCTGGGCTTAGAGATTTGGATCC	690
QY	661	GGG	CGGAGATACCGGCTCTGTGACGTGATCTGGGGGCTCGAGGTCTGCTCCGGC	720
Db	691	GGG	CGGAGATACCGGCTCTGTGACGTGATCTGGGGGCTCGAGGTCTGCTCCGGC	750
QY	721	TGG	AGACACACATGTGGAAGGCGACCGTTTCCCTCGAGATTCACGTGGTCACTCAGC	780
Db	751	TGG	AGACACACATGTGGAAGGCGACCGTTTCCCTCGAGATTCACGTGGTCACTCAGC	810
QY	781	ACC	GCTTTGACAGAGTTGACAGAGGCTTGGGGGCGCCGGAGAGCTGGCGTGTGGCC	840
Db	811	ACC	GCTTTGACAGAGTTGACAGAGGCTTGGGGGCGCCGGAGAGCTGGCGTGTGGCC	870
QY	841	GCT	TTCTGGAAGAGGGCCCGGAGAAACAGTGCCTATAGAGATTTGCTGCTGCTG	900
Db	871	GCT	TTCTGGAAGAGGGCCCGGAGAAACAGTGCCTATAGAGATTTGCTGCTGCTG	930
QY	901	GAA	GAATCTGCTGAGAGAGGCTCAGAGACTCAGSTCCACAGCTGACATATCTGCATC	960
Db	931	GAA	GAATCTGCTGAGAGAGGCTCAGAGACTCAGSTCCACAGCTGACATATCTGCATC	990
QY	961	CTG	CCCTGACATTCAGCGGCTACTTCAATATAGAGGATCTCTGACTACACCGCCCTGT	1020
Db	991	CTG	CCCTGACATTCAGCGGCTACTTCAATATAGAGGATCTCTGACTACACCGCCCTGT	1050
QY	1021	GCC	AGGGGTCTATCTGACCTGTGTTTAACAGACAGTATGCTGAGTGTAGACATC	1080
Db	1051	GCC	AGGGGTCTATCTGACCTGTGTTTAACAGACAGTATGCTGAGTGTAGACATC	1110
QY	1081	CAC	ACCCTCTGACACCCCTGGGGGACCTGTGTACTCTGCGGCTACAGCTGAACCTCCGA	1140
Db	1111	CAC	ACCCTCTCTGACACCCCTGGGGGACCTGTGTACTCTGCGGCTACAGCTGAACCTCCGA	1170
QY	1141	GCG	ACGACAGCTTTGAAATGAGCGAGTGAATGAGGCTCTTCCCTGCTGAAGTGAACAC	1200
Db	1171	GCG	ACGACGACGCTTTGAAATGAGCGAGTGAATGAGGCTCTTCCCTGCTGAAGTGAACAC	1230
QY	1201	AGT	CTCGGCTCTGACCAATGACAGCTGAATTCGCTGTGCTCTGTGTATACATCTTA	1260
Db	1231	AGT	CTCGGCTCTGACCAATGACAGCTGAATTCGCTGTGCTCTGTGTATACATCTTA	1290
QY	1261	GCC	CTGATTTTGGCTCTTTTCTGTACACAGGCTGGCGTTCCTTGTGACAAATGAGA	1320
Db	1291	GCC	CTGATTTTGGCTCTTTTCTGTACACAGGCTGGCGTTCCTTGTGACAAATGAGA	1350

QY	1322	AGCAGCCACAGAGGGGGACCAAGAGGGGGGTGTAGCTTACCGCCACGACAGAGGTATACGAG	1380
QY <td>1351</td> <td>AGCAGCACAGAGGGGACCAAGAGGGGGGTGTAGCTTACCGCCACGACAGAGGTATACGAG</td> <td>1410</td>	1351	AGCAGCACAGAGGGGACCAAGAGGGGGGTGTAGCTTACCGCCACGACAGAGGTATACGAG	1410
Db <td>1381</td> <td>ACTGAGCCTTAAGAGCTTGATCTTGGAGAAATGTGAGAACCCAGACAGAGGCATCTGAGGG</td> <td>1440</td>	1381	ACTGAGCCTTAAGAGCTTGATCTTGGAGAAATGTGAGAACCCAGACAGAGGCATCTGAGGG	1440
QY <td>1411</td> <td>ACTGAGCCTTAAGAGCTTGATCTTGGAGAAATGTGAGAACCCAGACAGAGGCATCTGAGGG</td> <td>1470</td>	1411	ACTGAGCCTTAAGAGCTTGATCTTGGAGAAATGTGAGAACCCAGACAGAGGCATCTGAGGG	1470
Db <td>1441</td> <td>GGAGCCGGTAACCTGTCTGTCTCTCATTTATGCGCACTTCCTTTTAACTGCGCAAGAAATT</td> <td>1500</td>	1441	GGAGCCGGTAACCTGTCTGTCTCTCATTTATGCGCACTTCCTTTTAACTGCGCAAGAAATT	1500
QY <td>1471</td> <td>GGAGCCGGTAACCTGTCTGTCTCTCATTTATGCGCACTTCCTTTTAACTGCGCAAGAAATT</td> <td>1530</td>	1471	GGAGCCGGTAACCTGTCTGTCTCTCATTTATGCGCACTTCCTTTTAACTGCGCAAGAAATT	1530
Db <td>1501</td> <td>TTTTAAATTAATTAATTTTAAAT</td> <td>1522</td>	1501	TTTTAAATTAATTAATTTTAAAT	1522
QY <td>1531</td> <td>TTTTAAATTAATTAATTTTAAAT</td> <td>1552</td>	1531	TTTTAAATTAATTAATTTTAAAT	1552

RESULT 11
ADY76124

ID	ABX76124	standard; DNA; 1552 BP.
----	----------	-------------------------

AC ABX76124;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polynucleotide #1.

KM Lung cancer-associated polynucleotide, gene, cell, cytostatic, emphysema;
KM antituberculous; antituberculous; non-small cell lung cancer; adenocarcinoma;
KM small cell lung cancer; benign lesion, precancerous lesion; bronchitis;
KM chronic obstructive pulmonary disease, hypersensitivity pneumonitis;
KM interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

Unidentified.

PN WO200286443-A2

PD 31-OCT-2002.

PF 18-APR-2002; 2002WO-US012476.

PR 18-APR-2001; 2001US-0284770P.

PR 09-NOV-2001; 2001US-0339245P.

PR 29-NOV-2001; 2001US-0334370P.

XX

[illegible]

XX	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
----	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

P-PSDB; ABU56408.

PT Detecting a lung cancer-associated

PT patient with a polynucleotide

EXPT 08901 TM TUNG CUNEE.
XX

XX

transcript in a cell from a E

to a sequence that is at least

associated polynucleotides are

inhibiting proliferation of a

administering a modulatory co

CC For treating lung cancer, such

lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the invention

Sequence 1552 BP; 302 A; 471 C; 461 G; 318 T; 0 U; 0 Other;

Query Match 100.0%; Score 1522; DB 7; Length 1552;

Best Local Similarity 100.0%; Pident No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACATGACGCGCATGAGTCCCTGTCCTCCAGCCCTGAGCTCCCTGTGATCCCGCC 60
 Db 31 ACATGACGCGCATGAGTCCCTGTCCTCCAGCCCTGAGCTCCCTGTGATCCCGCC 90
 QY 61 CCGCTCCAGGCTCACTGTGCACTGTGCTGCACTGTGCTTCTGATGCTGTCAT 120
 Db 91 CCGCTCCAGGCTCACTGTGCACTGTGCTGCACTGTGCTTCTGATGCTGTCAT 150
 QY 121 CCCAGAGGTGCCCCGAGTGCAGAGAGATTCCCTTGGAGAGAGCTTTCGGGGA 180
 Db 151 CCCAGAGGTGCCCCGAGTGCAGAGAGATTCCCTTGGAGAGAGCTTTCGGGGA 210
 QY 181 GATGACCACTGAGGAGAGATCTGCCAGTGAAGAGATTCCAGAGAGAGAT 240
 Db 211 GATGACCACTGAGGAGAGATCTGCCAGTGAAGAGATTCCAGAGAGAGAT 270
 QY 241 CCACCGGAGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGAGATCTACT 300
 Db 271 CCACCGGAGAGAGATCTACCTGAGAGAGATCTACCTGAGAGAGAGATCTACT 330
 QY 301 GAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGAAGTTAAGAGATCTACTACT 360
 Db 331 GAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGAAGTTAAGAGATCTACTACT 390
 QY 361 GAGGCTCTGAGAGATCTCAAGAACCCCAAGTAATGCCCCAGAGAGAGAGAGAT 420
 Db 391 GAGGCTCTGAGAGATCTCAAGAACCCCAAGTAATGCCCCAGAGAGAGAGAGAT 450
 QY 421 GACAGAGTCACTGAGGCTATGAGAGGAGCCGCGCTGAGCCCGGAGTCCCAAGCTGC 480
 Db 451 GACAGAGTCACTGAGGCTATGAGAGGAGCCGCGCTGAGCCCGGAGTCCCAAGCTGC 510
 QY 481 GCGGCGCGCTTCCAGTCCCGGTGATTCGCGCCCAAGCTGCGCTTTCGCGGCG 540
 Db 511 GCGGCGCGCTTCCAGTCCCGGTGATTCGCGCCCAAGCTGCGCTTTCGCGGCG 570
 QY 541 CTGCGCCCTGGAACTCTGGGCTTCCAGCTCCCGCGCTCCCAAGACTGCGCTGCG 600
 Db 571 CTGCGCCCTGGAACTCTGGGCTTCCAGCTCCCGCGCTCCCAAGACTGCGCTGCG 630
 QY 601 AACATGCGCACTGTGCACTGACCTGCTCTGTGGCTGAGAGATGCTGTGAGTCC 660
 Db 631 AACATGCGCACTGTGCACTGACCTGCTCTGTGGCTGAGAGATGCTGTGAGTCC 690
 QY 661 GGGCGGAGAGTACCGGGCTGTGAGTGTGCACTGCACTGGGGGCTGCAAGTGTGCGGG 720
 Db 691 GGGCGGAGAGTACCGGGCTGTGAGTGTGCACTGCACTGGGGGCTGCAAGTGTGCGGG 750
 QY 721 TCGAGCACACTGTGAGAGGCAACGCTTCCCTGCGAGATCCAGTGTTCACCTGAG 780
 Db 751 TCGAGCACACTGTGAGAGGCAACGCTTCCCTGCGAGATCCAGTGTTCACCTGAG 810
 QY 781 ACCGCTTTGCCAGATGACAGAGGCTTGGGGGCGCCGAGAGGCTGCGGAGTGTGCG 840
 Db 811 ACCGCTTTGCCAGATGACAGAGGCTTGGGGGCGCCGAGAGGCTGCGGAGTGTGCG 870
 QY 841 GCCTTTCTGAGAGAGGCGCGAAGAAACAGTGTCTATGACATTTGCTGTGCTG 900

Db 871 GCCTTTCTGAGAGAGGCGCGAAGAAACAGTGTCTATGAGCAGTGTCTGTGCTTG 930
 QY 901 GAGAAATCGCTGAGAGAGGCTCAGAGCTAGAGTCCCAAGACTGACATATCTGCACTC 960
 Db 931 GAGAAATCGCTGAGAGAGGCTCAGAGCTAGAGTCCCAAGACTGACATATCTGCACTC 990
 QY 961 CTGCGCTGTGACTGAGCGGCTACTTCCATATAGAGGGTCTGTGACTACCGCCCTGT 1020
 Db 991 CTGCGCTGTGACTGAGCGGCTACTTCCATATAGAGGGTCTGTGACTACCGCCCTGT 1050
 QY 1021 GCCCAGGTGTCTATGAGTGTGCTTTTAAACAGCACTGATGCTGATGCTAAGCAGCTC 1080
 Db 1051 GCCCAGGTGTCTATGAGTGTGCTTTTAAACAGCACTGATGCTGATGCTAAGCAGCTC 1110
 QY 1081 CACACCTCTCTGACACCCCTGTGGGAGACCTGTGATCTGTGGCTACAGCTGAACCTTCGA 1140
 Db 1111 CACACCTCTCTGACACCCCTGTGGGAGACCTGTGATCTGTGGCTACAGCTGAACCTTCGA 1170
 QY 1141 GCGAGCAGCCTTTGAATGAGGAGATGATGAGGCTCTTCCCTGTGAGTGTGACAGC 1200
 Db 1171 GCGAGCAGCCTTTGAATGAGGAGATGATGAGGCTCTTCCCTGTGAGTGTGACAGC 1230
 QY 1201 AGTCTCGAGCTGTGAGCAGCTCAGCTGATTCCTGCTGAGCTGTGATCCTTA 1260
 Db 1231 AGTCTCGAGCTGTGAGCAGCTCAGCTGATTCCTGCTGAGCTGTGATCCTTA 1290
 QY 1261 GCCCTGTTTTTGGCTCTCTTTTGTCTCAACAGCGTCCGCTTCTTTGTGACATGAGA 1320
 Db 1291 GCCCTGTTTTTGGCTCTCTTTTGTCTCAACAGCGTCCGCTTCTTTGTGACATGAGA 1350
 QY 1321 AGCAGCAGAGAGAGGAGCAACAAAGGGGTGTGATACCGCCAGAGAGATGACCGAG 1380
 Db 1351 AGCAGCAGAGAGAGGAGCAACAAAGGGGTGTGATACCGCCAGAGAGATGACCGAG 1410
 QY 1381 ACTGAGCCTTAGAGCTGATCTTGAAGATGTGAAGAGCCAGCAGAGCATCTGAGG 1440
 Db 1411 ACTGAGCCTTAGAGCTGATCTTGAAGATGTGAAGAGCCAGCAGAGCATCTGAGG 1470
 QY 1441 GAGGCGGTAACCTGCTCTGCTGCTCATTAATGCACTTCTTTTAACTGCCAAGAAAT 1500
 Db 1471 GAGGCGGTAACCTGCTCTGCTGCTCATTAATGCACTTCTTTTAACTGCCAAGAAAT 1530
 QY 1501 TTTTAAATAAATATTTAATAAT 1522
 Db 1531 TTTTAAATAAATATTTAATAAT 1552

RESULT 12
 AAH26551
 ID AAH26551 standard; cDNA, 1833 BP.
 XX
 AC AAH26551;
 XX
 DT 12-NOV-2001 (first entry)
 XX
 DE DNA encoding G250-GM-CSF fusion protein.
 XX
 XX GM-CSF; granulocyte-macrophage colony stimulating factor; G250; vaccine;
 XX genetic immunisation; adoptive immunotherapy; therapy; kidney cancer;
 XX antigen; renal cell carcinoma; tumour; human; gene therapy; antitumour;
 XX ss.
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX Key
 FT CDS Location/Qualifiers
 FT 1..1815
 FT /tag= a
 FT /product= "G250-GM-CSF"
 FT /transl_except= (pos:433..435, aa:Arg)
 FT /transl_except= (pos:436..438, aa:Arg)
 XX
 PN WO200160317-A2.

XX MN CDNA clone.
 DE
 XX MN; endogenous; Matu; quasi-viral agent; human; mammary tumour;
 KW classical virus; slow virus; prion; exogenous MX; p58X;
 KW cytoplasmic antigen; conservative; Hela cell; twin protein; p54/58N;
 KW cell surface; nucleus; monoclonal antibody; Mab W75; neoplasm;
 KW pre-neoplastic disease; vaccine; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 1..1290
 FT CDS /*tag= a
 FT
 XX MO9318152-A1.
 PN
 XX 16-SEP-1993.
 PD
 XX 08-MAR-1993; 93WO-US002024.
 PF
 XX 11-MAR-1992; 92CS-00000709.
 PR 21-OCT-1992; 92US-00964589.
 XX
 PA (CIBA) CIBA CORNING DIAGNOSTICS CORP.
 PA (VIRO-) INST VIRROLOGY.
 PI Zavada J, Pastorekova S, Pastorek J;
 XX WPI: 1993-303466/38.
 DR P-PsDB; AAR41746.
 XX
 PT New MN gene and polypeptide(s) - used in diagnosis, prognosis and therapy
 of neoplastic and/or pre-neoplastic disease.
 XX
 PS Claim 2; Fig 1; 72pp; English.
 XX
 CC This sequence represents the intronless MN gene which is a cellular gene
 CC which is the endogenous component of the Matu agent. Matu is a novel
 CC quasi-viral agent with rather unusual properties. It is presumably
 CC derived from a human mammary tumour. In some aspects it resembles
 CC classical viruses, whereas in other respects it resembles "slow" viruses of
 CC (prions), and in still other aspects it is different from both classes of
 CC viruses. Matu is a two component system. One part of the complex, which
 CC exogenous MX, is transmissible, and is manifest by a protein, p58X, which
 CC is a cytoplasmic antigen which reacts with some natural sera of humans
 CC and of various animals. The other component, MN, is endogenous to human
 CC cells. MN is a cellular gene showing very little homology with known DNA
 CC sequences. It is rather conservative and present as a single copy in the
 CC chromosomal DNA of various vertebrates. MN is manifest in Hela cells by a
 CC twin protein p54/58N, that is localised on the cell surface and in the
 CC nucleus. Immunoblot using a monoclonal antibody reactive with p54/58N
 CC (Mab W75) reveals two bands at 54 kD and 58 kD. These two bands may
 CC correspond to one type of protein that differs by glycosylation pattern
 CC or by how it is processed. The expression of the MN gene is strongly
 CC correlated with tumourigenicity. MN products can be used in/can be used
 CC in diagnostic and/or prognostic assays for neoplastic and/or pre-
 CC neoplastic disease. MN polypeptides, produced recombinantly by
 CC unicellular hosts, can also be used for antibody production and in
 CC vaccines for inducing protective immunity against neoplastic disease and
 CC a dampening effect upon tumourigenic activity. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 CC
 XX
 SQ Sequence 1397 BP; 284 A; 402 C; 425 G; 286 T; 0 U; 0 Other;
 Query Match 85.6%; Score 1302.8; DB 2; Length 1397;
 Best Local Similarity 99.3%; Pred. No. 1.2e-307;
 Matches 1392; Conservative 0; Mismatches 2; Indels 8; Gaps 8;
 QY 124 CAGAGTTGCCCGGATGCAAGAGATTCCCTTGGAGAGAGGCTCTTCTGGGAGAT 183
 Db 1 CAGAGTTGCCCGGATGCAAGAGATTCCCTTGGAGAGAGGCTCTTCTGGGAGAT 59

QY 184 GACCACTGGCGAGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGATCCA 243
 Db 60 GACCACTGGCGAGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGATCCA 119
 QY 244 CCCGAGAGAGAGATCTCACTGGAGAGAGAGATCTCACTGGAGAGAGATCTCACTGAA 303
 Db 120 CCCGAGAGAGAGATCTCACTGGAGAGAGAGATCTCACTGGAGAGAGATCTCACTGAA 179
 QY 304 GTTAA-GCCTAATCAGAGAGAGAGGCTCCCTGAAATTGAGAGATCTCACTGTTGA 362
 Db 180 GTTAAATGCTAATCAGAGAGAGAGGCTCCCTGAAATTGAGAGATCTCACTGTTGA 239
 QY 353 GGCCTCGAGATCTCAAGAACCCAGATTAATGCCCCAGAGACAAAGAGAGATGA 422
 Db 240 GGCCTCGAGATCTCAAGAACCCAGATTAATGCCCCAGAGACAAAGAGAGATGA 299
 QY 423 CCAGATCATTTGGCGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
 Db 300 CCAGATCATTTGGCGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358
 QY 483 GGGCCGCTTCCAGTCCCGGTGATATCCGCCCCAGCTGCGCCCTTCCGCGGCGCT 542
 Db 359 GGGCCGCTTCCAGTCCCGGTGATATCCGCCCCAGCTGCGCCCTTCCGCGGCGCT 418
 QY 543 GGGCCGCTTCCAGTCCCGGTGATATCCGCCCCAGCTGCGCCCTTCCGCGGCGCA 602
 Db 419 GGGCCGCTTCCAGTCCCGGTGATATCCGCCCCAGCTGCGCCCTTCCGCGGCGCA 477
 QY 603 CAATGGCCAGATGTCAGTCACTGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTG 662
 Db 478 CAATGGCCAGATGTCAGTCACTGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTG 537
 QY 663 GCGGAGTACCGGGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
 Db 538 GCGGAGTACCGGGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
 QY 723 GGAGCACTGTGAG 782
 Db 597 GGAGCACTGTGAG 656
 QY 783 GGCCTTTCAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842
 Db 657 GGCCTTTCAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 715
 QY 843 CTTTGTGAG 901
 Db 716 CTTTGTGAG 775
 QY 902 AAGAAATGCTGAG 961
 Db 776 AAGAAATGCTGAG 835
 QY 962 TCCCTCTGACTTACGCGCTTCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
 Db 836 TCCCTCTGACTTACGCGCTTCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAG 895
 QY 1022 CCCAGAGTCTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
 Db 896 CCCAGAGTCTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 955
 QY 1082 ACACCTCTCTGACACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
 Db 956 ACACCTCTCTGACACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1015
 QY 1142 CGAGCAGAGCTTTGAATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201
 Db 1016 CGAGCAGAGCTTTGAATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1075
 QY 1202 GTCTCTGGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1261
 Db 1076 GTCTCTGGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1135
 QY 1262 CCTGTGTTTGGCCCTCTTTTGTCTGCACAGAGTGGGCTTCTGTGAGATGAGAA 1321

Accession	Sequence	Position
Db	1136 CCTGGATTTTGGGCGCTTTTGTGTGCACACAGCTGGGCTTCCTGTGGAGATGAGGA	1195
Qy	1322 GGAGAGACGAAGAGGGAAACAAAGGGGTGTGAGC-TACGCGCCACACAGAGTACCAG	1380
Db	1196 GGCAGACAGAAAGGGAAACAAAGGGGTGTGAGCGTACCGCCACAGAGGTAGCCAG	1255
Qy	1381 ACTGAGCCTTAAGAGCTGGATCTTGGAGATGTGAGAGCCAGCCAGAGGCATCTGAGG	1440
Db	1256 ACTGGAGCCTTAAGAGGTGGATCTTGGAGAAATGTAGAGCCAGCCAGAGGCATCTGAGG	1315
Qy	1441 GGAGCCGATACGTGCTGTCCTGTCTATTGTCACATTCCTTTAACTGCCAAGAAAT	1500
Db	1316 GGAGCCGATACGTGCTGTCCTGTCTATTGTCACATTCCTTTAACTGCCAAGAAAT	1375
Qy	1501 TTTTAAATTAATATTATTAAAT 1522	
Db	1376 TTTTAAATTAATATTATTAAAT 1397	

[illegible]

```

FT      /*tag= 9
FT      /number= 6
FT      exon      8776..8933
FT      /*tag= h
FT      /number= 7
FT      exon      9446..9590
FT      /*tag= i
FT      /number= 8
FT      exon      9705..9731
FT      /*tag= j
FT      /number= 9
FT      exon      10349..10430
FT      /*tag= k
FT      /number= 10
FT      exon      10561..10751
FT      /*tag= l
FT      /number= 11
XX      W09534650-A2.
XX      PD      21-DEC-1995.
XX      PF      15-JUN-1995; 95WO-US007628.
XX      PR      15-JUN-1994; 94US-00260190.
XX      PR      07-JUN-1995; 95US-00477504.
XX      PR      07-JUN-1995; 95US-00481658.
XX      PR      07-JUN-1995; 95US-00485049.
XX      PR      07-JUN-1995; 95US-00485862.
XX      PR      07-JUN-1995; 95US-00485863.
XX      PR      07-JUN-1995; 95US-00486756.
XX      PR      07-JUN-1995; 95US-00487077.
XX      PA      (CIBA ) CIBA CORNING DIAGNOSTICS CORP.
XX      PA      (VIRO-) INST VIROLOGY.
XX      PI      Zavada J, Pastorekova S, Pastorek J;
XX      DR      WPI; 1996-049679/05.
XX      PT      MN gene, protein and nucleic acid fragments - used as primers and probes
XX      PT      in the detection of MN antigens and antibodies, and in the treatment of
XX      PT      (pre)neoplastic disease.
XX      PS      Claim 1; Fig 3; 102pp; English.
XX      CC      The present sequence is the complete Mutu endogenous cellular component,
XX      CC      MN, genomic DNA, which was isolated from lymphocytic choriomeningitis
XX      CC      virus (LCMV) infected HeLa cells. Persistent LCMV, the exogenous Mutu
XX      CC      transmissible agent (MX), infection increases the expression level of the
XX      CC      MN gene. MN is a putative oncogene, and can therefore be used in the
XX      CC      development of prods. for the diagnosis and treatment of neoplastic (NP),
XX      CC      or pre-NP diseases. NP diseases can be treated using DNA antisense to MN
XX      CC      transcribed mRNA, anti-MN protein antibodies can be used for the
XX      CC      diagnosis NP or pre-NP diseases and a vaccine contg. immunogenic amounts
XX      CC      of the MN protein can be used to immunise a vertebrate against a NP
XX      CC      disease associated with MN antigen expression. (Updated on 25-MAR-2003 to
XX      CC      correct PR field.)
XX      SQ      Sequence 10897 BP; 2654 A; 2738 C; 2647 G; 2857 T; 0 U; 1 Other.
XX
Query Match      27.2%; Score 414.4; DB 2; Length 10897;
Best Local Similarity 99.8%; Pred. No. 1.4e-90;
Matches 415; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 ACAGTCAGCCGATGCTCCCTGTCAGCCAGCCCTGCTCCCTGTCATCCGAGCC 60
Db      3536 ACAGTCAGCCGATGCTCCCTGTCAGCCAGCCCTGCTCCCTGTCATCCGAGCC 3595
QY      61 CTTGCTCAGGCTCACTGTCAGTCTGTCATGTCATGTCATGTCATGTCATGTCAT 120
Db      3596 CTTGCTCAGGCTCACTGTCAGTCTGTCATGTCATGTCATGTCATGTCATGTCAT 3655

```

```

QY      121 CCCGAGGTTGCCCGGATGACAGAGGATTCCTCCCTTGAGAGAGGCTCTTGAGGAA 180
Db      3656 CCCGAGGTTGCCCGGATGACAGAGGATTCCTCCCTTGAGAGAGGCTCTTGAGGAA 3715
QY      181 GATGACCACTGGGCGAGAGGATCTGCCAGTCGAGAGGATTCACCCAGAGAGGAGAT 240
Db      3716 GATGACCACTGGGCGAGAGGATCTGCCAGTCGAGAGGATTCACCCAGAGAGGAGAT 3775
QY      241 CCACCCGAGAGAGGATCTACCTGAGAGGAGATCTACCTGAGAGAGGATCTACCT 300
Db      3776 CCACCCGAGAGAGGATCTACCTGAGAGGAGATCTACCTGAGAGAGGATCTACCT 3835
QY      301 GAAGTTAAGCTTAATCAGAGAAAGAGGCTCCCTGAGATTAGAGATCTACTGTT 360
Db      3836 GAAGTTAAGCTTAATCAGAGAAAGAGGCTCCCTGAGATTAGAGATCTACTGTT 3895
QY      361 GAGGCTCCTGAGATCCCTCAAGAACCCAGAAATATGCCACAGAGGCAAGAAAG 416
Db      3896 GAGGCTCCTGAGATCCCTCAAGAACCCAGAAATATGCCACAGAGGCAAGAAAG 3951

```

Search completed: March 7, 2004, 03:28:27
Job time : 472 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 03:11:59 ; Search time 96 Seconds
(without alignments)
8798.284 Million cell updates/sec

Title: US-09-967-237-1

Perfect score: 1522
Sequence: 1 acagtcagcgcagcagtcctc.....ttaataatattataat 1522

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1522	100.0	1522	2	US-08-481-658B-1
2	1522	100.0	1522	2	US-08-477-504A-1
3	1522	100.0	1522	2	US-08-486-756A-1
4	1522	100.0	1522	2	US-08-485-862B-1
5	1522	100.0	1522	3	US-08-787-739-1
6	1522	100.0	1522	3	US-08-487-077A-1
7	1522	100.0	1522	3	US-08-485-863A-1
8	1522	100.0	1522	3	US-08-485-049D-1
9	1522	100.0	1522	3	US-09-178-115-1
10	1522	100.0	1522	3	US-09-177-776-1
11	1399	91.9	1399	3	US-08-335-469-1
12	1302.8	85.6	1397	5	US-07-964-589-1
13	1302.8	85.6	1397	5	PCT-US93-02024-1
14	414.4	27.2	10898	2	US-08-481-658B-5
15	414.4	27.2	10898	2	US-08-477-504A-5
16	414.4	27.2	10898	2	US-08-486-756A-5
17	414.4	27.2	10898	2	US-08-485-862B-5
18	414.4	27.2	10898	2	US-08-787-739-5
19	414.4	27.2	10898	3	US-08-487-077A-5
20	414.4	27.2	10898	3	US-08-485-863A-5
21	414.4	27.2	10898	3	US-08-485-049D-5
22	414.4	27.2	10898	3	US-09-178-115-5
23	414.4	27.2	10898	3	US-09-177-776-5
24	413.4	27.2	415	2	US-08-481-658B-28
25	413.4	27.2	415	2	US-08-477-504A-28
26	413.4	27.2	415	2	US-08-486-756A-28
27	413.4	27.2	415	2	US-08-485-862B-28

28	413.4	27.2	415	3	US-08-487-077A-28	Sequence 28, Appl
29	413.4	27.2	415	3	US-08-485-863A-28	Sequence 28, Appl
30	413.4	27.2	415	3	US-08-485-049D-28	Sequence 28, Appl
31	413.4	27.2	445	3	US-08-787-739-28	Sequence 28, Appl
32	413.4	27.2	445	3	US-09-178-115-28	Sequence 28, Appl
33	413.4	27.2	445	3	US-09-177-776-28	Sequence 28, Appl
34	233.4	15.3	470	2	US-08-481-658B-55	Sequence 55, Appl
35	233.4	15.3	470	2	US-08-477-504A-55	Sequence 55, Appl
36	233.4	15.3	470	2	US-08-486-756A-55	Sequence 55, Appl
37	233.4	15.3	470	2	US-08-485-862B-55	Sequence 55, Appl
38	233.4	15.3	470	3	US-08-787-739-55	Sequence 55, Appl
39	233.4	15.3	470	3	US-08-487-077A-55	Sequence 55, Appl
40	233.4	15.3	470	3	US-08-485-863A-55	Sequence 55, Appl
41	233.4	15.3	470	3	US-08-485-049D-55	Sequence 55, Appl
42	233.4	15.3	470	3	US-09-178-115-55	Sequence 55, Appl
43	233.4	15.3	470	3	US-09-177-776-55	Sequence 55, Appl
44	191	12.5	191	2	US-08-481-658B-38	Sequence 38, Appl
45	191	12.5	191	2	US-08-477-504A-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-1
Sequence 1, Application US/08481658B
Patent No. 5955075
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jarmilr
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1522 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
Query Match 100.0%; Score 1522; DB 2; Length 1522;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-477-504A-1

Query Match 100.0%; Score 1522; DB 2; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGTCAGCCGATGAGCTCCCTGTCGCCCCAGCCCTGCTCCCTGTTGATCCGAGCC 60
 DB 1 ACAGTCAGCCGATGAGCTCCCTGTCGCCCCAGCCCTGCTCCCTGTTGATCCGAGCC 60
 QY 61 CCTGCTCCAGGCTCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 DB 61 CCTGCTCCAGGCTCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 QY 121 CCCAGAGGTTGCCCCGAGTGAAGAGATTCCTCCCTTGGAGAGAGCTCTTCTGGAGAA 180
 DB 121 CCCAGAGGTTGCCCCGAGTGAAGAGATTCCTCCCTTGGAGAGAGCTCTTCTGGAGAA 180
 QY 181 GATGACCCATGCGGCGAGAGAGATTCCTCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
 DB 181 GATGACCCATGCGGCGAGAGAGATTCCTCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
 QY 241 CCACCCGAGAGAGAGATTCACCTGAGAGAGAGATTCACCTGAGAGAGAGATTCACCT 300
 DB 241 CCACCCGAGAGAGAGATTCACCTGAGAGAGAGATTCACCTGAGAGAGAGATTCACCT 300
 QY 301 GAAGTTAAGCTTAATCAGAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 DB 301 GAAGTTAAGCTTAATCAGAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 361 GAGGCTCTGAGAGATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 DB 361 GAGGCTCTGAGAGATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 421 GACAGAGATCATTTGGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 DB 421 GACAGAGATCATTTGGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 481 GCGGAGCGCTTCCAGTCCCGGCTGATTCGCGCCCCAGCTGCGCGCTTTCGCGCGCC 540
 DB 481 GCGGAGCGCTTCCAGTCCCGGCTGATTCGCGCCCCAGCTGCGCGCTTTCGCGCGCC 540
 QY 541 CTGCGCCCTGGAATCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGCGCTGCGC 600
 DB 541 CTGCGCCCTGGAATCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGCGCTGCGC 600
 QY 601 AACATAGGCACTGCTGCACTGACCTGCTCTGAGGCTGAGATGAGATGAGATGAG 660
 DB 601 AACATAGGCACTGCTGCACTGACCTGCTCTGAGGCTGAGATGAGATGAGATGAG 660
 QY 661 GGGGAG 720
 DB 661 GGGGAG 720
 QY 721 TCGAGACACATGCTGAG 780
 DB 721 TCGAGACACATGCTGAG 780
 QY 781 ACCGCTTTGCGAG 840
 DB 781 ACCGCTTTGCGAG 840
 QY 841 GCCTTTCTGAG 900
 DB 841 GCCTTTCTGAG 900
 QY 901 GAAGAATGCTGAG 960
 DB 901 GAAGAATGCTGAG 960

DB 901 GAAGAATGCTGAG 960
 QY 961 CTGCGCTTGAATTCAGCGGCTTACTTCAATATAGAGAGAGAGAGAGAGAGAGAGAG 1020
 DB 961 CTGCGCTTGAATTCAGCGGCTTACTTCAATATAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 1021 GCCCAG 1080
 DB 1021 GCCCAG 1080
 QY 1081 CACACCTTCTGACACCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 DB 1081 CACACCTTCTGACACCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 QY 1141 GCGAG 1200
 DB 1141 GCGAG 1200
 QY 1201 AGCTCTGAG 1260
 DB 1201 AGCTCTGAG 1260
 QY 1261 GCGCTGAG 1320
 DB 1261 GCGCTGAG 1320
 QY 1321 AGCGAG 1380
 DB 1321 AGCGAG 1380
 QY 1381 ACTGAG 1440
 DB 1381 ACTGAG 1440
 QY 1441 GAG 1500
 DB 1441 GAG 1500
 QY 1501 TTTTAAATTAATTTTAAAT 1522
 DB 1501 TTTTAAATTAATTTTAAAT 1522

RESULT 3
 US-08-486-756A-1
 Sequence 1, Application US/08486756A
 Patent No. 5981711
 GENERAL INFORMATION:
 APPLICANT: Zavada, Jan
 APPLICANT: Pastorekova, Silvia
 APPLICANT: Pastorek, Jaromir
 TITLE OF INVENTION: MN Gene and Protein
 NUMBER OF SEQUENCES: 86
 CORRESPONDENCE ADDRESSES:
 ADDRESSES: Leona L. Lauder
 STREET: 6 Mariposa Court
 CITY: Tiburon
 STATE: California
 COUNTRY: USA
 ZIP: 94920
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/260,190
 FILING DATE: 15-JUN-1994
 ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
 REGISTRATION NUMBER: 30,863
 REFERENCE/DOCKET NUMBER: D-0021.3C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-435-2034
 TELEFAX: 415-435-0727
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1522 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-486-756A-1

Query Match 100.0%; Score 1522; DB 2; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAGTACGCGGATGAGTCCCTGTCGCGCCGAGCCCTGAGTCCCTGTTGATCCGCGCC 60
 1 AAGTACGCGGATGAGTCCCTGTCGCGCCGAGCCCTGAGTCCCTGTTGATCCGCGCC 60
 61 CCTGCTCAGGCTCAGTGTGCACTGTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 61 CCTGCTCAGGCTCAGTGTGCACTGTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 121 CCCGAGAGTTGCGCGGATGAGAGGATTTCCCTTGGAGAGAGGCTCTTTCGAGGAA 180
 121 CCCGAGAGTTGCGCGGATGAGAGGATTTCCCTTGGAGAGAGGCTCTTTCGAGGAA 180
 181 GATGACCACTGCGGAGAGAGATCTGCGCAGTGAAGAGATTACCCAGAGAGAGAT 240
 181 GATGACCACTGCGGAGAGAGATCTGCGCAGTGAAGAGATTACCCAGAGAGAGAT 240
 241 CACCCGAGAGAGAGATCTACTGGAAGAGAGATCTACTGGAAGAGAGATCTACTGGA 300
 241 CACCCGAGAGAGAGATCTACTGGAAGAGAGATCTACTGGAAGAGAGATCTACTGGA 300
 241 CACCCGAGAGAGAGATCTACTGGAAGAGAGATCTACTGGAAGAGAGATCTACTGGA 300
 301 GAAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAAGTTAGAGATCTACTGCTT 360
 301 GAAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAAGTTAGAGATCTACTGCTT 360
 361 GAGGCTCTGAGATCTCTAAGAACCCAGAAATATGCCCAAGAGAGAGAGAGAT 420
 361 GAGGCTCTGAGATCTCTAAGAACCCAGAAATATGCCCAAGAGAGAGAGAT 420
 421 GACCAAGATCATTTGGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
 421 GACCAAGATCATTTGGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
 481 GCGGCGCTTCCAGTCCCGGTGATATCCGCCCCAGCTGCGCCCTTCTGCGCGCC 540
 481 GCGGCGCTTCCAGTCCCGGTGATATCCGCCCCAGCTGCGCCCTTCTGCGCGCC 540
 541 CTGCGCCCTTGAATCTCTGAGCTTCCAGCTCCCGCGCTCCAGAACTGCGCTCGC 600
 541 CTGCGCCCTTGAATCTCTGAGCTTCCAGCTCCCGCGCTCCAGAACTGCGCTCGC 600
 601 AACATGCGCAAGTGTGCACTGACCTGCTCTGAGCTGAGAGATGCTCTGAGTCCC 660
 601 AACATGCGCAAGTGTGCACTGACCTGCTCTGAGCTGAGAGATGCTCTGAGTCCC 660
 661 GCGGCGGAGTACCGGCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 720
 661 GCGGCGGAGTACCGGCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 720
 721 TCGGAGCACACTGTGAG 780
 721 TCGGAGCACACTGTGAG 780

781 ACCGCTTTGCGAGAGTTCAGAGGCTTTGGGCGCCCGGAGAGGCTGAGCGTGTGCGC 840
 781 ACCGCTTTGCGAGAGTTCAGAGGCTTTGGGCGCCCGGAGAGGCTGAGCGTGTGCGC 840
 841 GCGCTTTGAGAGAGGCGCCGAGAGAAACAGTCTATGAGAGAGTGTCTGCTGCTG 900
 841 GCGCTTTGAGAGAGGCGCCGAGAGAAACAGTCTATGAGAGAGTGTCTGCTGCTG 900
 901 GAAGAAATCCCTGAGAGAGGCTCAGAGACTCAGTGTCCAGAGACTGAGACATCTGACTC 960
 901 GAAGAAATCCCTGAGAGAGGCTCAGAGACTCAGTGTCCAGAGACTGAGACATCTGACTC 960
 961 CTGCGCTCTGACTGAGCGGCTACTTCCAAATATGAGGGTCTCTGACTACACCGCCCT 1020
 961 CTGCGCTCTGACTGAGCGGCTACTTCCAAATATGAGGGTCTCTGACTACACCGCCCT 1020
 1021 GCGGAGGAGTCAATCTGAGAGTGTGTTAACCCAGAGAGTATGAGTGTGAGTCTGAGAGCTC 1080
 1021 GCGGAGGAGTCAATCTGAGAGTGTGTTAACCCAGAGAGTATGAGTGTGAGTCTGAGAGCTC 1080
 1081 CACACCTCTCTGACACCTGTGAGGAGACCTGAGTCTTGGCTACAGCTGAACCTTCCGA 1140
 1081 CACACCTCTCTGACACCTGTGAGGAGACCTGAGTCTTGGCTACAGCTGAACCTTCCGA 1140
 1141 GCGAGCAGCTTTGAAATGAGGAGATGATGAGGCTCTTCCCTGCTGAGTGAAGAGC 1200
 1141 GCGAGCAGCTTTGAAATGAGGAGATGATGAGGCTCTTCCCTGCTGAGTGAAGAGC 1200
 1201 AGTCTGCGGCTGTGAGAGCAGTCACTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1201 AGTCTGCGGCTGTGAGAGCAGTCACTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1261 GCGCTGCTTTTGGCTCTCTTTTGGCTGTCACAGAGGCTGCTGCTGCTGCTGCTGCTGCT 1320
 1261 GCGCTGCTTTTGGCTCTCTTTTGGCTGTCACAGAGGCTGCTGCTGCTGCTGCTGCTGCT 1320
 1321 GCGCTGCTTTTGGCTCTCTTTTGGCTGTCACAGAGGCTGCTGCTGCTGCTGCTGCTGCT 1380
 1321 GCGCTGCTTTTGGCTCTCTTTTGGCTGTCACAGAGGCTGCTGCTGCTGCTGCTGCTGCT 1380
 1381 ACTGAGCTTGAAGGCTGATCTTGAAGATGTGAAGAGCCAGAGAGAGATCTGAGAGG 1440
 1381 ACTGAGCTTGAAGGCTGATCTTGAAGATGTGAAGAGCCAGAGAGAGATCTGAGAGG 1440
 1441 GAGCGGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 1441 GAGCGGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 1501 TTTTAAATTAATTTTAAAT 1522
 1501 TTTTAAATTAATTTTAAAT 1522

RESULT 4
 US-08-485-862B-1
 Sequence 1, Application US/08485862B
 Patent No. 5989838
 GENERAL INFORMATION:
 APPLICANT: Zavada, Jan
 APPLICANT: Pastorekova, Silvia
 APPLICANT: Pastorek, Jaromir
 TITLE OF INVENTION: NY Gene and Protein
 NUMBER OF SEQUENCES: 86
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Leona L. Lauder
 STREET: 6 Mariposa Court
 CITY: Tiburon
 STATE: California
 COUNTRY: USA
 ZIP: 94920
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,862B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/477,504
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/260,190
 FILING DATE: 15-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: lauder, leona L.
 REGISTRATION NUMBER: 30,863
 REFERENCE/DOCKET NUMBER: D-0021.3D
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-435-2034
 TELEFAX: 415-435-0727
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1522 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-485-862B-1

Query Match 100.0%; Score 1522; DB 2; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGTACGCGGAGTGGCTCCCTGTCGCGCCGAGCCCTGCTCCCTGTTGATCCGCGCC 60
 DB 1 ACAGTACGCGGAGTGGCTCCCTGTCGCGCCGAGCCCTGCTCCCTGTTGATCCGCGCC 60
 QY 61 CCTGCTCCAGGCTCCACGTCGCACTGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 DB 61 CCTGCTCCAGGCTCCACGTCGCACTGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 QY 121 CCCGAGAGTGGCCCGGAGTCAGAGAGATTCCTGCTGAGAGAGGCTTTCTGAGGAA 180
 DB 121 CCCGAGAGTGGCCCGGAGTCAGAGAGATTCCTGCTGAGAGAGGCTTTCTGAGGAA 180
 QY 121 CCCGAGAGTGGCCCGGAGTCAGAGAGATTCCTGCTGAGAGAGGCTTTCTGAGGAA 180
 DB 121 CCCGAGAGTGGCCCGGAGTCAGAGAGATTCCTGCTGAGAGAGGCTTTCTGAGGAA 180
 QY 181 GATGACCACTGGGCGAGAGAGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
 DB 181 GATGACCACTGGGCGAGAGAGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
 QY 181 GATGACCACTGGGCGAGAGAGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
 DB 181 GATGACCACTGGGCGAGAGAGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
 QY 241 CCACCCGAGAGAGATCTACTGAGAGAGATCTACTGAGAGAGATCTACTGAGAGAGATCTACT 300
 DB 241 CCACCCGAGAGAGATCTACTGAGAGAGATCTACTGAGAGAGATCTACTGAGAGAGATCTACT 300
 QY 301 GAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAAGTTAAGATCTACTGATTT 360
 DB 301 GAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAAGTTAAGATCTACTGATTT 360
 QY 301 GAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAAGTTAAGATCTACTGATTT 360
 DB 301 GAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAAGTTAAGATCTACTGATTT 360
 QY 361 GAGGCTCTGAGATCTCCAGAGAGAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAT 420
 DB 361 GAGGCTCTGAGATCTCCAGAGAGAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAT 420
 QY 361 GAGGCTCTGAGATCTCCAGAGAGAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAT 420
 DB 361 GAGGCTCTGAGATCTCCAGAGAGAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAT 420
 QY 421 GACGAGAGTCAATGAGGAG 480
 DB 421 GACGAGAGTCAATGAGGAG 480
 QY 481 GAGGAGCGCTTCAGTCCCGAGTGAATCCGCGCCGAGCTGCGCGCTTCTGCGCGCC 540
 DB 481 GAGGAGCGCTTCAGTCCCGAGTGAATCCGCGCCGAGCTGCGCGCTTCTGCGCGCC 540
 QY 541 CTGCGCGCTTCAGTCCCGAGTGAATCCGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCGCG 600
 DB 541 CTGCGCGCTTCAGTCCCGAGTGAATCCGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCGCG 600

QY 601 AACATGCGCACAGTGTGCACTGACCTGCTCCCTGAGCTGAGAGATGAGTCTGAGGTC 660
 DB 601 AACATGCGCACAGTGTGCACTGACCTGCTCCCTGAGCTGAGAGATGAGTCTGAGGTC 660
 QY 661 GGGCGGAGATCCGGGCTCTGCAAGTGTGCACTGAGTGGGGGCTGAGAGTCTGCGGGC 720
 DB 661 GGGCGGAGATCCGGGCTCTGCAAGTGTGCACTGAGTGGGGGCTGAGAGTCTGCGGGC 720
 QY 661 GGGCGGAGATCCGGGCTCTGCAAGTGTGCACTGAGTGGGGGCTGAGAGTCTGCGGGC 720
 DB 661 GGGCGGAGATCCGGGCTCTGCAAGTGTGCACTGAGTGGGGGCTGAGAGTCTGCGGGC 720
 QY 721 TCGAGACACATGTGAG 780
 DB 721 TCGAGACACATGTGAG 780
 QY 721 TCGAGACACATGTGAG 780
 DB 721 TCGAGACACATGTGAG 780
 QY 781 ACCGCTTTGCGAGAGTGAACAGAGCTTGGGGGCGCGGAGAGAGAGAGAGAGAGAGAG 840
 DB 781 ACCGCTTTGCGAGAGTGAACAGAGCTTGGGGGCGCGGAGAGAGAGAGAGAGAGAGAG 840
 QY 781 ACCGCTTTGCGAGAGTGAACAGAGCTTGGGGGCGCGGAGAGAGAGAGAGAGAGAGAG 840
 DB 781 ACCGCTTTGCGAGAGTGAACAGAGCTTGGGGGCGCGGAGAGAGAGAGAGAGAGAGAG 840
 QY 841 GCCTTCTGAG 900
 DB 841 GCCTTCTGAG 900
 QY 841 GCCTTCTGAG 900
 DB 841 GCCTTCTGAG 900
 QY 901 GAGAAATGCTGAG 960
 DB 901 GAGAAATGCTGAG 960
 QY 901 GAGAAATGCTGAG 960
 DB 901 GAGAAATGCTGAG 960
 QY 961 CTGCGCTTCTGAG 1020
 DB 961 CTGCGCTTCTGAG 1020
 QY 961 CTGCGCTTCTGAG 1020
 DB 961 CTGCGCTTCTGAG 1020
 QY 1021 GCCAGAGGTGTCATCTGAG 1080
 DB 1021 GCCAGAGGTGTCATCTGAG 1080
 QY 1021 GCCAGAGGTGTCATCTGAG 1080
 DB 1021 GCCAGAGGTGTCATCTGAG 1080
 QY 1081 CACACCTCTGTCAGACAG 1140
 DB 1081 CACACCTCTGTCAGACAG 1140
 QY 1081 CACACCTCTGTCAGACAG 1140
 DB 1081 CACACCTCTGTCAGACAG 1140
 QY 1141 GCGACGAGAGCTTTGAG 1200
 DB 1141 GCGACGAGAGCTTTGAG 1200
 QY 1141 GCGACGAGAGCTTTGAG 1200
 DB 1141 GCGACGAGAGCTTTGAG 1200
 QY 1201 AGTCTGAGAGCTCTGAG 1260
 DB 1201 AGTCTGAGAGCTCTGAG 1260
 QY 1201 AGTCTGAGAGCTCTGAG 1260
 DB 1201 AGTCTGAGAGCTCTGAG 1260
 QY 1261 GCCCTGATTTTGGCTCTCTTTTGGCTCTCTTTTGGCTCTCTTTTGGCTCTCTTTTGGCT 1320
 DB 1261 GCCCTGATTTTGGCTCTCTTTTGGCTCTCTTTTGGCTCTCTTTTGGCTCTCTTTTGGCT 1320
 QY 1261 GCCCTGATTTTGGCTCTCTTTTGGCTCTCTTTTGGCTCTCTTTTGGCTCTCTTTTGGCT 1320
 DB 1261 GCCCTGATTTTGGCTCTCTTTTGGCTCTCTTTTGGCTCTCTTTTGGCTCTCTTTTGGCT 1320
 QY 1321 AGGACGACAG 1380
 DB 1321 AGGACGACAG 1380
 QY 1321 AGGACGACAG 1380
 DB 1321 AGGACGACAG 1380
 QY 1381 ACTGAGAGCTGAG 1440
 DB 1381 ACTGAGAGCTGAG 1440
 QY 1381 ACTGAGAGCTGAG 1440
 DB 1381 ACTGAGAGCTGAG 1440
 QY 1441 GAGGCGGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 DB 1441 GAGGCGGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 1501 TTTTAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1522
 DB 1501 TTTTAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1522

RESULT 5
 US-08-787-739-1
 Sequence 1, Application US/08787739
 Patent No. 6027887
 GENERAL INFORMATION:
 APPLICANT: Zavada, Jan
 APPLICANT: Pastorekova, Silvia
 APPLICANT: Pastorek, Jaromir

TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1522 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ANTI-SENSE: NO
US-08-787-739-1
Query Match 100.0%; Score 1522; DB 3; Length 1522;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGTCAGCCGATGCTGCTCCCTGTCGCCAGCCCTGCTCCCTCTTTATATCCGGCC 60
DB 1 ACAGTCAGCCGATGCTGCTCCCTGTCGCCAGCCCTGCTCCCTCTTTATATCCGGCC 60
QY 61 CCGCTCCAGGCTCTCACTGTCGAACCTGCTGTCACCTGCTCTTCTGATCCCTGCAT 120
DB 61 CCGCTCCAGGCTCTCACTGTCGAACCTGCTGTCACCTGCTCTTCTGATCCCTGCAT 120
QY 121 CCCCAGAGGTTGCCCCGGATGCAAGAGATTCCCTTGGAGAGAGGCTCTTTGGGGAA 180
DB 121 CCCCAGAGGTTGCCCCGGATGCAAGAGATTCCCTTGGAGAGAGGCTCTTTGGGGAA 180
QY 181 GATGACCACTGCGGAGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGGAT 240

DB 181 GATGACCACTGCGGAGAGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGAT 240
QY 241 CCACCCGAGAGAGAGATCTTACCTGAGAGAGAGATTTACCTGAGAGAGATCTTACT 300
DB 241 CCACCCGAGAGAGAGATCTTACCTGAGAGAGAGATTTACCTGAGAGAGATCTTACT 300
QY 301 GAAGTTAAGCTTAATCAAGAGAGAGAGGCTCCCTGAAGTTAGAGATCTTACTACTGT 360
DB 301 GAAGTTAAGCTTAATCAAGAGAGAGAGGCTCCCTGAAGTTAGAGATCTTACTACTGT 360
QY 361 GAGGCTCTGAGATCTTCAAGAAACCCAGAAATATGCCACAGGAGCAAAAGAGAGAT 420
DB 361 GAGGCTCTGAGATCTTCAAGAAACCCAGAAATATGCCACAGGAGCAAAAGAGAGAT 420
QY 421 GACCAAGATCAATGGCGGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 GACCAAGATCAATGGCGGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 GCGGAGCGCTTCAAGTCCCGAGTGAATATCCGCCCCAGCTGCGCGCTTCTGCGCGCC 540
DB 481 GCGGAGCGCTTCAAGTCCCGAGTGAATATCCGCCCCAGCTGCGCGCTTCTGCGCGCC 540
QY 541 CTGGGCGCGCTTCAAGTCCCGAGTGAATATCCGCCCCAGCTGCGCGCTTCTGCGCGCC 600
DB 541 CTGGGCGCGCTTCAAGTCCCGAGTGAATATCCGCCCCAGCTGCGCGCTTCTGCGCGCC 600
QY 601 AACATGAGCAAGTGTGCAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 AACATGAGCAAGTGTGCAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GGGGCGGAGTACCGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GGGGCGGAGTACCGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 TCGAGAGCACTGAG 780
DB 721 TCGAGAGCACTGAG 780
QY 781 ACCGCTTTCGCAAGTGTGCAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 781 ACCGCTTTCGCAAGTGTGCAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 GCGCTTCTGAG 900
DB 841 GCGCTTCTGAG 900
QY 901 GAAAGATGCTGAG 960
DB 901 GAAAGATGCTGAG 960
QY 961 CTGGCTGCTGAG 1020
DB 961 CTGGCTGCTGAG 1020
QY 1021 GCGGAGGCTGATCTGAG 1080
DB 1021 GCGGAGGCTGATCTGAG 1080
QY 1081 CAGACCTCTGAG 1140
DB 1081 CAGACCTCTGAG 1140
QY 1141 GCGAGAGAGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 GCGAGAGAGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 AGTCTGAG 1260
DB 1201 AGTCTGAG 1260
QY 1261 GCGCTGATTTTGGGCTCTTCTTCTGCTGCAAGAGCTGCGCTTCTTCTGAGATGAG 1320

QY 841 GCCTTCTGGAGAGAGCCCGGAGAGAAACAGTGTCTATGACAGTTCCTCTCTGCTTG 900
DB 841 GCCTTCTGGAGAGAGCCCGGAGAGAAACAGTGTCTATGACAGTTCCTCTCTGCTTG 900
QY 901 GAGAAATGCTGAGAGAGGCTCAGAGACTCAGGTCCAGGACTGACATATCTGCACTC 960
DB 901 GAGAAATGCTGAGAGAGGCTCAGAGACTCAGGTCCAGGACTGACATATCTGCACTC 960
QY 961 CTGGCTCTGACTTCAAGCCGCTACTTCCATATGAGGGGTCTCTGACTAACCAGCCCTGT 1020
DB 961 CTGGCTCTGACTTCAAGCCGCTACTTCCATATGAGGGGTCTCTGACTAACCAGCCCTGT 1020
QY 1021 GCCCAGGGTGTCTATCTGAGCTGTGTTTAACCAAGACAGTGTCTGATGCTAAGACCTC 1080
DB 1021 GCCCAGGGTGTCTATCTGAGCTGTGTTTAACCAAGACAGTGTCTGATGCTAAGACCTC 1080
QY 1081 CACACCTCTCTGACACCCCTGTGGGACCTTGCTGACTCTCGGCTACAGCTGAACCTTCCGA 1140
DB 1081 CACACCTCTCTGACACCCCTGTGGGACCTTGCTGACTCTCGGCTACAGCTGAACCTTCCGA 1140
QY 1141 GCGACGACACCTTTGATGAGGAGATGATGAGGCTCTCTCTGCTGAGAGTGAACAGC 1200
DB 1141 GCGACGACACCTTTGATGAGGAGATGATGAGGCTCTCTCTGCTGAGAGTGAACAGC 1200
QY 1201 AGTCTCGGGCTGTGAGCCAGTCCAGCTGAATTCCTGCTGCTGTGTGAACATCTTA 1260
DB 1201 AGTCTCGGGCTGTGAGCCAGTCCAGCTGAATTCCTGCTGCTGTGTGAACATCTTA 1260
QY 1261 GCCCTGTGTTTGGGCTCTCTTTTGTCTGACACAGCGGTGGGTCTCTGCTGAGATGAG 1320
DB 1261 GCCCTGTGTTTGGGCTCTCTTTTGTCTGACACAGCGGTGGGTCTCTGCTGAGATGAG 1320
QY 1321 AGCGACACAGAGAGGAGAACCAAGAGGGGTGTGAGCTACCGCCAGACAGAGTACCGAG 1380
DB 1321 AGCGACACAGAGAGGAGAACCAAGAGGGGTGTGAGCTACCGCCAGACAGAGTACCGAG 1380
QY 1381 ACTGAGCCTTAAGGCTGTGATCTTGGAGAAATGTAAGAACGACAGAGGATCTAGAGG 1440
DB 1381 ACTGAGCCTTAAGGCTGTGATCTTGGAGAAATGTAAGAACGACAGAGGATCTAGAGG 1440
QY 1441 GGAGCCGGTAAGTCTCTGTCTCTGCTCATTAAGCACTTCCTTTAACTGCCAAGAAATT 1500
DB 1441 GGAGCCGGTAAGTCTCTGTCTCTGCTCATTAAGCACTTCCTTTAACTGCCAAGAAATT 1500
QY 1501 TTTTAAATAAATAATTATAAT 1522
DB 1501 TTTTAAATAAATAATTATAAT 1522

RESULT 9
US-09-178-115-1
; Sequence 1, Application US/09178115
; Patent No. 6297041
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021, SA
; CURRENT APPLICATION NUMBER: US/09/178, 115
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 09/177, 776
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787, 739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485, 049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486, 756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477, 504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481, 658
; EARLIER FILING DATE: 1995-06-07

EARLIER APPLICATION NUMBER: 08/485, 862
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485, 863
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/487, 077
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/260, 190
EARLIER FILING DATE: 1994-06-15
EARLIER APPLICATION NUMBER: 08/177, 093
EARLIER FILING DATE: 1993-12-30
EARLIER APPLICATION NUMBER: 07/964, 589
EARLIER FILING DATE: 1992-10-21
EARLIER APPLICATION NUMBER: PV-709-92
EARLIER FILING DATE: 1992-03-11
NUMBER OF SEQ ID NOS: 116
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 1
LENGTH: 1522
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: CDS
LOCATION: (13)..(1389)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (124)..(1389)
US-09-178-115-1

Query Match 100.0%; Score 1522; DB 3; Length 1522;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGTCAAGCCGATGCTCCCTGTGTCGCCAGCCCTGCTCTCTGTGATCCCGGCT 60
DB 1 ACAGTCAAGCCGATGCTCCCTGTGTCGCCAGCCCTGCTCTCTGTGATCCCGGCT 60
QY 61 CTTGCTCAGGCTCTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 CTTGCTCAGGCTCTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 CCCCAGAGTTCGCCGAGATGACAGAGATTCCTCTGAGAGAGCTCTTCTGGGAA 180
DB 121 CCCCAGAGTTCGCCGAGATGACAGAGATTCCTCTGAGAGAGCTCTTCTGGGAA 180
QY 181 GATGACCACTGGGCGAGAGAGATTCGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
DB 181 GATGACCACTGGGCGAGAGAGATTCGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
QY 241 CCACCCGAGAGAGATCTTACTGAGAGAGATCTTACTGAGAGAGATCTTACT 300
DB 241 CCACCCGAGAGAGATCTTACTGAGAGAGATCTTACTGAGAGAGATCTTACT 300
QY 301 GAACTTAAGCTTAATCAGAGAGAGGCTCCCTGGAATGAGATTAAGTACTTACTGTT 360
DB 301 GAACTTAAGCTTAATCAGAGAGAGGCTCCCTGGAATGAGATTAAGTACTTACTGTT 360
QY 361 GAGGCTCTGAGATCTCTCAGAACCCAGAAATATGCCACAGGAGCAAGAGAGGAT 420
DB 361 GAGGCTCTGAGATCTCTCAGAACCCAGAAATATGCCACAGGAGCAAGAGAGGAT 420
QY 421 GACCAAGTATGAGGCTATGAGAGGACCCGCTGACCCCGGCTGCTCCCAAGCTGC 480
DB 421 GACCAAGTATGAGGCTATGAGAGGACCCGCTGACCCCGGCTGCTCCCAAGCTGC 480
QY 481 GCGGCGGCTTCAGTCCCGGATGATCCGCCCCCACTGCGGCTTCTGCGCGGCT 540
DB 481 GCGGCGGCTTCAGTCCCGGATGATCCGCCCCCACTGCGGCTTCTGCGCGGCT 540
QY 541 CTGGGCGGCTTCAGTCCCGGCTTCCAGCTTCCGCGCTCCAGAACTGGGCTGGGC 600
DB 541 CTGGGCGGCTTCAGTCCCGGCTTCCAGCTTCCGCGCTCCAGAACTGGGCTGGGC 600
QY 601 AACAAATGAGCAAGTGTGCACTGACCTGCTCTGAGCTGAGATGAGCTGAGTCCC 660

421 GACCAAGATATGCGCTATGAGGAGACCCGCTTGGCCCCGGGTGTCCCAAGCTTGC 480
421 GACCAAGATATGCGCTATGAGGAGACCCGCTTGGCCCCGGGTGTCCCAAGCTTGC 480
481 GCGGGCGGCTTCCAGTCCCGGATGATATCCGCCCCAGCTGGCGCTTGTGCGCGG 540
481 GCGGGCGGCTTCCAGTCCCGGATGATATCCGCCCCAGCTGGCGCTTGTGCGCGG 540
541 CTGGCCCCCTGGAATCTCTGGGCTTCCAGCTTCCGCGCTTCCAGACTGGCGCTGCG 600
541 CTGGCCCCCTGGAATCTCTGGGCTTCCAGCTTCCGCGCTTCCAGACTGGCGCTGCG 600
601 AACATGGCCACATGATGCACTGACCTGCTCTGGGTAGATGCTTGGAGTCCC 660
601 AACATGGCCACATGATGCACTGACCTGCTCTGGGTAGATGCTTGGAGTCCC 660
661 GGGCGGAGTACCGGGGCTTGGACCTGATCTGACCTGGGGGCTGACAGTGTCCGGCC 720
661 GGGCGGAGTACCGGGGCTTGGACCTGATCTGACCTGGGGGCTGACAGTGTCCGGCC 720
721 TCGAGACACATGATGAGAGGACCCGCTTCCGCGAGATCCAGTGGTTCACCTCAGC 780
721 TCGAGACACATGATGAGAGGACCCGCTTCCGCGAGATCCAGTGGTTCACCTCAGC 780
781 ACCGCTTTGGCAGAGTTGACGAGGCTTGGGGGCGCCGGGAGGCTTGGCTGTGGCC 840
781 ACCGCTTTGGCAGAGTTGACGAGGCTTGGGGGCGCCGGGAGGCTTGGCTGTGGCC 840
841 GCGTTCTGAGAGAGGCGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
841 GCGTTCTGAGAGAGGCGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
901 GAAGAAATCGCTGAG 960
901 GAAGAAATCGCTGAG 960
961 CTGGCTTCTGACTTACAGCGCTTACTTCCATATGAGGGGCTTCTGACTACAGCGCTGT 1020
961 CTGGCTTCTGACTTACAGCGCTTACTTCCATATGAGGGGCTTCTGACTACAGCGCTGT 1020
1021 GCCAGAGGTGCTCATCTGAGCTGTATTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1021 GCCAGAGGTGCTCATCTGAGCTGTATTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1081 CACACCTCTCTGACACCTGTGTGGGAGCTGTGACTCTTGGCTTACAGCTTGAATCCGA 1140
1081 CACACCTCTCTGACACCTGTGTGGGAGCTGTGACTCTTGGCTTACAGCTTGAATCCGA 1140
1141 GCGAGGAGAGCTTGAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1141 GCGAGGAGAGCTTGAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1201 AGTCTCGGGGCTGCTGAGCAGTCCAGCTGAATCTCTGCTGCTGCTGCTGCTGCTGCT 1260
1201 AGTCTCGGGGCTGCTGAGCAGTCCAGCTGAATCTCTGCTGCTGCTGCTGCTGCTGCT 1260
1261 GCGCTGCTTGTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
1261 GCGCTGCTTGTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
1321 AGGCAAGCAG 1380
1321 AGGCAAGCAG 1380
1381 ACTGAGGCTGAG 1440
1381 ACTGAGGCTGAG 1440
1441 GAGAGCGGTAACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1500
1441 GAGAGCGGTAACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1500
1501 TTTTAAATTAATTTTAAAT 1522

Db 1501 TTTTAAATTAATTTTAAAT 1522

RESULT 11
US-08-335-469-1
Sequence 1, Application US/08335469A
Patent No. 604535
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: Methods to Detect and Quantify MN Protein/Polypeptide
TITLE OF INVENTION: Using MN-specific Antibodies
FILE REFERENCE: D-0021A
CURRENT APPLICATION NUMBER: US/08/335,469A
CURRENT FILING DATE: 1994-11-07
EARLIER APPLICATION NUMBER: 07/964,589
EARLIER FILING DATE: 1992-10-21
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1399
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1266)
US-08-335-469-1

Query Match 91.9%; Score 1399; DB 3; Length 1399;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

124 CAGAGGTGCCCCGAGATGAGAGAGATTCCTTGGAGAGAGAGAGAGAGAGATTC 183
1 CAGAGGTGCCCCGAGATGAGAGAGATTCCTTGGAGAGAGAGAGAGAGATTC 60
184 GACCACTGGGCGAGAGAGATTCGCCAGTGAAGAGATTCACCCAGAGAGAGATTC 243
61 GACCACTGGGCGAGAGAGATTCGCCAGTGAAGAGATTCACCCAGAGAGAGATTC 120
244 CCGGAGAGAGAGATTCCTGAGAGAGAGATTCACCTGAGAGAGAGAGATTC 303
121 CCGGAGAGAGAGATTCCTGAGAGAGAGATTCACCTGAGAGAGAGAGATTC 180
304 GTTAAGCTTAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
181 GTTAAGCTTAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
364 GTCCTGAGAGATTCCTCAAGATCCCAAGATTAAGCCCAAGAGAGAGAGAGAG 423
241 GTCCTGAGAGATTCCTCAAGATCCCAAGATTAAGCCCAAGAGAGAGAGAGAG 300
424 CAGAGTATTGGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
301 CAGAGTATTGGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
484 GGGCGCTTCAAGTCCCGGAGATATCCGCCCCAGAGAGAGAGAGAGAGAGAGAG 543
361 GGGCGCTTCAAGTCCCGGAGATATCCGCCCCAGAGAGAGAGAGAGAGAGAGAG 420
544 CCGCCCCCTGGAATCTCTGGGCTTCCAGCTTCCGCGCTTCCAGAACTGCGCTGCG 603
421 CCGCCCCCTGGAATCTCTGGGCTTCCAGCTTCCGCGCTTCCAGAACTGCGCTGCG 480
604 AATGGCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
481 AATGGCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
664 CCGGAGTACCGGAGCTTGGAGCTGAGATCTGCACTGGAGAGAGAGAGAGAGAG 723
541 CCGGAGTACCGGAGCTTGGAGCTGAGATCTGCACTGGAGAGAGAGAGAGAGAG 600

Mon Mar 8 10:16:30 2004

us-09-967-237-1.rtf

Page 14

Qy	663	GCGGAGTACCGGAGCTCTGCAAGTCACTCACTGAGGGAGCTGCAGATCTGTCGGGCTC	722
Dp	538	GGCGAGTACCCGCTCTGCAAGTCTGCACTGCACTGGGGGGCTGCAGATCTGTCGGGCTC	596
Qy	723	GGAGCACTCTGTGGAAGGCGACCGTTTCCCTGCCGAACTCCAGTGTCTACCTCAGCAC	782
Dp	597	GGAGCACTCTGTGGAAGGCGACCGTTTCCCTGCCGAACTCCAGTGTCTACCTCAGCAC	656
Qy	783	CGCCTTTGCGAGAGTTGACGAGGCGCTTGGGGCGCCCGGAGAGGCGTGGCGTTGGCGGC	842
Dp	657	CGCCTTTGCGAGAGTTGACGAGGCGCTTGGGGCGCCCGGAGAGGCGTGGCGTTGGG - GGC	715
Qy	843	CTTTCTGAGAGAGGGCCCGGAAAGAAACAGG - CCTGTAGACAGTTGCTGTCCGTTGG	901
Dp	716	CTTTCTGAGAGAGGGCCCGGAAAGAAACAGGTCCTATAGACAGTTGCTGTCCGTTGG	775
Qy	902	AAGAAATCGCTAGAGGAAGCTCAGAGCTCAGAGTCCGAGAGCTGGAACATATCTGCATCC	961
Dp	776	AAGAAATCGCTAGAGGAAGCTCAGAGCTCAGAGTCCGAGAGCTGGAACATATCTGCATCC	835
Qy	962	TGGCCCTGACTTCAAGCCGACTCACTTCCAAATAGAGGGGCTCTGACATACCGCCCTGG	1021
Dp	836	TGGCCCTGACTTCAAGCCGACTCACTTCCAAATAGAGGGGCTCTGACATACCGCCCTGG	895
Qy	1022	CCGAGGAGTGCATCTGGACTGTGTTAAACGACAGTATGCTAGTGTCAAGCAAGCTCC	1081
Dp	896	CCGAGGAGTGCATCTGGACTGTGTTAAACGACAGTATGCTAGTGTCAAGCAAGCTCC	955
Qy	1082	ACACCGCTCTGACACCCGTGAGGAGCTGTGTACTCTCGAGTACAGCTGAATCCCGAG	1141
Dp	956	ACACCGCTCTGACACCCGTGAGGAGCTGTGTACTCTCGAGTACAGCTGAATCCCGAG	1015
Qy	1142	CGAGCGAGCTTTGAAATGGGCGAGTGAATTGAGGCTCTCCCTTCCCTGCTGAGTGAACGCA	1201
Dp	1016	CGAGCGAGCTTTGAAATGGGCGAGTGAATTGAGGCGCTCTCCCTGCTGAGTGAACGCA	1075
Qy	1202	GTCCTCGAGGCTGTGAGCAGTCCACAGCTGAATTCCTGCTGAGCTGTGTAATCTCAG	1261
Dp	1076	GTCCTCGAGGCTGTGAGCAGTCCACAGCTGAATTCCTGCTGAGCTGTGTAATCTCAG	1135
Qy	1262	CCCTGGTTTTTGGCCTCCTTTTGTCTGTACCAAGCGTCCGCTTCTTGTGCAATGAGAA	1321
Dp	1136	CCCTGGTTTTTGGCCTCCTTTTGTCTGTACCAAGCGTCCGCTTCTTGTGCAATGAGAA	1195
Qy	1322	GGCAGACAGAAAGGGGAACCAAGAGGGGTGTGAGC - TACCGCCACAGAGATGACGAG	1380
Dp	1196	GGCAGACAGAAAGGGGAACCAAGAGGGGTGTGAGC - TACCGCCACAGAGATGACGAG	1255
Qy	1381	ACTGGAAGCTTAAGAGCTGTGATCTTGGAAATGTGAGAAAGCCACACAGAGCATCTGAGGG	1440
Dp	1256	ACTGGAAGCTTAAGAGAGCTGTGATCTTGGAAATGTGAGAAAGCCACACAGAGCATCTGAGGG	1315
Qy	1441	GGAGCGGATACGTCTCTGCTCTGCTCAATTATGCCACTTCTTTTAAGTCCAGAAATT	1500
Dp	1316	GGAGCGGATACGTCTCTGCTCTGCTCAATTATGCCACTTCTTTTAAGTCCAGAAATT	1375
Qy	1501	TTTTAAATTAATATTTTAAAT 1522	
Dp	1376	TTTTAAATTAATATTTTAAAT 1397	

RESULT 13
PCT-US93-02024-1
: Sequence 1, Application PC/TUS9302024

APPLICANT: CIBA Corning Diagnostics Corp.
APPLICANT: Institute of Virology
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona J. Lander
STREET: Steuart Street Tower, 18th Fl., One Market
STREET: Plaza

CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94105
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/02024
 FILING DATE: 19930308
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: CS FV-709-92
 FILING DATE: 10-MAR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/964,589
 FILING DATE: 21-OCT-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: lauder, leona L
 REGISTRATION NUMBER: 30,863
 REFERENCE/DOCKET NUMBER: D-0021
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-777-9257
 TELEFAX: 415-543-4219
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1397 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1287
 PCT-US93-02024-1

Query Match	85.6%	Score 1302.8	DB 5	Length 1397
Best Local Similarity	99.3%	Prod. No. 0		
Matches 1392	Conservative	0	Mismatches 2	Indels 8
QY	124	CAGAGTTGCCCCGGATCCAGAGAGATTC	CCCTTGAGAGAGGCTCTTTGGGAGAGAT	183
Db	1	CAGAGTTGCCCCGGATCCAGAGAGATTC	CCCCCTTGAGAGAGGCTCTTTGGGAGAGAT	59
QY	184	GACCCATCTGGGAGAGAGAGATCTGCCAGTGAAGAGAGATTC	CCCAAGAGAGAGATTC	243
Db	60	GACCCATCTGGGAGAGAGAGATCTGCCAGTGAAGAGAGATTC	CCCAAGAGAGAGATTC	119
QY	244	CCCGAGAGAGAGATCTACCTGGAGAGAGATCTACCTGGAGAGAGATTC	TACTTGA	303
Db	120	CCCGAGAGAGAGATCTACCTGGAGAGAGATCTACCTGGAGAGAGATTC	TACTTGA	179
QY	304	GTTAA-GCTTAATCAGAAAGAGGGCTCCGTAAGTGAAGAGATCTCAACGTGGA		362
Db	180	GTTAATGCTTAATCAGAAAGAGGGCTCCGTAAGTGAAGAGATCTCAACGTGGA		239
QY	363	GGCTCCCTGGAGATCTCACAAGACCCAGAAATAATGCCACAGGACAAAGAGGGAGTGA		422
Db	240	GGCTCCCTGGAGATCTCACAAGACCCAGAAATAATGCCACAGGACAAAGAGGGAGTGA		299
QY	423	CCAGATCATTTGGCGCTATGGAAGCGAACCGCCCTGGCCCCCGGGTGTCCCCAGCGTGGGC		482
Db	300	CCAGATCATTTGGCGCTATGGAAGCGAACCGC-CCTGGCCCCGGGTGTCCCCAGCGTGGGC		358
QY	483	GGCGCGCTTCCAGTCCCGGATGATATCGGCCCCCAAGCTGGCGCTTGTGCGCGGCTT		542
Db	359	GGCGCGCTTCCAGTCCCGGATGATATCGGCCCCCAAGCTGGCGCTTGTGCGCGGCTT		418
QY	543	GGCGCCCTCGAACTCTCGGCTTCCAGTCCCGCGCTCCCAAGACTGGGCTGGCA		602

419 GCGCCCTGGAATCTCTGGGCTTCCAGCTCCCGCGCTCCCGAACTGGCGCTGC-AGA 477
QY 603 CAATGGCCACAGTGTGCACTGACCTGCTCTGAGGCTAGAGATGGCTCTGGGTCGGG 662
Db 478 CAATGGCCACAGTGTGCACTGACCTGCTCTGAGGCTAGAGATGGCTCTGGGTCGGG 537
QY 663 GCGGAGTACCGGGCTCTGACAGCTGATCTGCACTGGGGGGGCTGAGAGTCTCCGGGCTC 722
Db 538 GCGGAGTACCGGGCTCTGACAGCTGATCTGCACTGGGGGGGCTGAGAGTCTCCGGGCTC 596
QY 723 GAGACACACTGTGGAAGGCCACCGTTTCCCTCCGAGATCAAGTGTTCACCTCAAGAC 782
Db 597 GAGACACACTGTGGAAGGCCACCGTTTCCCTCCGAGATCAAGTGTTCACCTCAAGAC 656
QY 783 GCGCTTGGCAGAGTTGACAGAGGCTTGGGGGGGCGGGGAGGCTGGCCGCTGTGGCGCC 842
Db 657 GCGCTTGGCAGAGTTGACAGAGGCTTGGGGGGGCGGGGAGGCTGGCCGCTGTGGCGCC 715
QY 843 CTTTCTGAGAGAGGGCCCGGAGAGAAACAAGTGC-CTATGAGCAGTTGCTGTCTGCTTGG 901
Db 716 CTTTCTGAGAGAGGGCCCGGAGAGAAACAAGTGTGCATGAGCAGTTGCTGTCTGCTTGG 775
QY 902 AAGAATGCTGAGAGAGAGGCTCAAGACTCAAGTCTCCAGAGCTGACATATCTGCACTTC 961
Db 776 AAGAATGCTGAGAGAGAGGCTCAAGACTCAAGTCTCCAGAGCTGACATATCTGCACTTC 835
QY 962 TGCCCTCTGACTGAGCGGCTACTTCCATATAGAGGGTCTCTGACTACAGCGCCCTGTG 1021
Db 836 TGCCCTCTGACTGAGCGGCTACTTCCATATAGAGGGTCTCTGACTACAGCGCCCTGTG 895
QY 1022 CCCAGGGTGTGATCTGAGCTGTGTTTAAACAGACAGTGTGCTGAGTCTTAAGCACTCC 1081
Db 896 CCCAGGGTGTGATCTGAGCTGTGTTTAAACAGACAGTGTGCTGAGTCTTAAGCACTCC 955
QY 1082 ACACGCTCTCTGACACCTGTGGGGAGCTGTGATCTGCGGCTGACAGCTGAACTTCCGAG 1141
Db 956 ACACGCTCTCTGACACCTGTGGGGAGCTGTGATCTGCGGCTGACAGCTGAACTTCCGAG 1015
QY 1142 CGACGACGCTTGAATGAGCGAGATGTTAGGCTCTCTCCCTGCTGAGTGAAGCA 1201
Db 1016 CGACGACGCTTGAATGAGCGAGATGTTAGGCTCTCTCCCTGCTGAGTGAAGCA 1075
QY 1202 GTCCTGGGGGCTGAGCAGCTCAGCTGAACTTCCGCGGCTGTGCTGAGTCACTCTGAG 1261
Db 1076 GTCCTGGGGGCTGAGCAGCTCAGCTGAACTTCCGCGGCTGTGCTGAGTCACTCTGAG 1135
QY 1262 CCGTGGTTTTGGGCTCTTTTGTCTGACACAGCGTGGGCTCTTGTGAGATGAGAA 1321
Db 1136 CCGTGGTTTTGGGCTCTTTTGTCTGACACAGCGTGGGCTCTTGTGAGATGAGAA 1195
QY 1322 GCGAGCACAAGAGGGGAAACCAAGAGGGGTGTAGGCT-ACCGGCCACAGAGGTACCGAG 1380
Db 1196 GCGAGCACAAGAGGGGAAACCAAGAGGGGTGTAGGCTACCGGCCACAGAGGTACCGAG 1255
QY 1381 ACTGAGCCTAGAGGCTGATCTTGTGAGATGTGAGAGCCAGCCAGAGGCACTGAGAG 1440
Db 1256 ACTGAGCCTAGAGGCTGATCTTGTGAGATGTGAGAGCCAGCCAGAGGCACTGAGAG 1315
QY 1441 GGAAGCCGCTAAGTCTCTGTCTGCTGCTATATGCTCCACTTCTTTTAACTGCAAGAA 1500
Db 1316 GGAAGCCGCTAAGTCTCTGTCTGCTGCTATATGCTCCACTTCTTTTAACTGCAAGAA 1375
QY 1501 TTTTAAATTAATTTTATAT 1522
Db 1376 TTTTAAATTAATTTTATAT 1397

RESULT 14
US-08-481-658B-5
; Sequence 5, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:

APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021,3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10898 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-5
Query Match 27.2%; Score 414.4; DB 2; Length 10898;
Best Local Similarity 99.8%; Pred. No. 7.4e-94;
Matches 415; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACAGTACCGCGATGAGCTCCCTGCTGCCCCAGCCCTTGCTCTGTTATCCGAGC 60
Db 3537 ACAGTACCGCGATGAGCTCCCTGCTGCCCCAGCCCTTGCTCTGTTATCCGAGC 3596
QY 61 CCGTCTCCAGGCTCACTGTGCACTGGCTGTGCACTGCTGTGATGAGCTGTGCAT 120
Db 3597 CCGTCTCCAGGCTCACTGTGCACTGGCTGTGCACTGCTGTGATGAGCTGTGCAT 3656
QY 121 CCCAGAGGTGCCCGGATGCAAGAGATTCCTTGGAGAGAGCTCTTGGGAA 180
Db 3657 CCCAGAGGTGCCCGGATGCAAGAGATTCCTTGGAGAGAGCTCTTGGGAA 3716
QY 181 GATGACCCACTGGGGAGAGAGATTCGCCAGGTGAAGAGATTACCCAGAGAGAGAT 240
Db 3717 GATGACCCACTGGGGAGAGAGATTCGCCAGGTGAAGAGATTACCCAGAGAGAGAT 3776
QY 241 CCACCCGAGAGAGAGATCTACCTGAGAGAGATCTTACTGAGAGAGATCTACT 300
Db 3777 CCACCCGAGAGAGAGATCTACCTGAGAGAGATCTTACTGAGAGAGATCTACT 3836
QY 301 GAAGTTAAGCTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 360
Db 3837 GAAGTTAAGCTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 3896
QY 361 GAGGCTCTGAGAGATCTCAAGAACCCAGATATATGCTCCACAGAGAGAGAGAG 416
Db 3897 GAGGCTCTGAGAGATCTCAAGAACCCAGATATATGCTCCACAGAGAGAGAGAG 3952

RESULT 15

US-08-477-504A-5
Sequence 5, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ. ID NO.: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10898 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-5

Query Match

27.2%; Score 414.4; DB 2; Length 10898;

Best Local Similarity

99.8%; Pred. No. 7,4e-94;

Matches 415; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ACAGTACGCGGCATGCTCCCTGTGCCCCCAAGCCCTGCTCTGTGATCCCGGCC	60
DB	3537	ACAGTACGCGGCATGCTCCCTGTGCCCCCAAGCCCTGCTCTGTGATCCCGGCC	3596
QY	61	CTGTCTCAGGCTCACTGTGCACTGTGCTGTCACTGTGCTTCTGATGCTGTCCAT	120
DB	3597	CTGTCTCAGGCTCACTGTGCACTGTGCTGTCACTGTGCTTCTGATGCTGTCCAT	3656
QY	121	CCCCAGAGTTGCCCCGGATGCAAGAGATTCCCCCTGGAGAGAGCTCTTTGGGGAA	180
DB	3657	CCCCAGAGTTGCCCCGGATGCAAGAGATTCCCCCTGGAGAGAGCTCTTTGGGGAA	3716
QY	181	GATGACCACTGGGCGAGAGAGATCTGCCAGTGAAGAGATTCAACCAGAGAGAGAT	240
DB	3717	GATGACCACTGGGCGAGAGAGATCTGCCAGTGAAGAGATTCAACCAGAGAGAGAT	3776
QY	241	CCACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATCTA	300
DB	3777	CCACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATCTA	3836

QY	301	GAAGTTAAGCCTTAATCAGAAAGAGAGGCTCCCTGAAAGTTAGAGATCTACTGTT	360
DB	3837	GAAGTTAAGCCTTAATCAGAAAGAGAGGCTCCCTGAAAGTTAGAGATCTACTGTT	3896
QY	361	GAGGCTCTGGAGATCTCTCAAGAACCCCAATATGCCCCACAGGACAAAGAGG	416
DB	3897	GAGGCTCTGGAGATCTCTCAAGAACCCCAATATGCCCCACAGGACAAAGAGG	3952

Search completed: March 7, 2004, 05:25:41
JOB time : 100 secs

; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1522 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; US-09-772-713-1

Query Match 100.0%; Score 1522; DB 9; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGTACGCGGATGGCTCCCTGTCGCGCCAGCCGCTTGGTGTGATCCGCGCC 60
 Db 1 AAGGTACCGGATGGCTCCCTGTCGCGCCAGCCGCTTGGTGTGATCCGCGCC 60
 QY 61 CCTGCTCCAGGCTTCACTGTGTGCACTGTGTGTCTGTCTGTGTGTGTGTGTGT 120
 Db 61 CCTGCTCCAGGCTTCACTGTGTGCACTGTGTGTCTGTCTGTGTGTGTGTGTGT 120
 QY 121 CCCGAGAGGTTGCCCCGATGAGAGAGATTCGCCCTTGGAGAGAGGCTTCTTGGGAA 180
 Db 121 CCCGAGAGGTTGCCCCGATGAGAGAGATTCGCCCTTGGAGAGAGGCTTCTTGGGAA 180
 QY 181 GATGACCCACTGGGCGAGAGAGATTCGCCCTTGGAGAGAGATTCGCCCTTGGGAA 240
 Db 181 GATGACCCACTGGGCGAGAGAGATTCGCCCTTGGAGAGAGATTCGCCCTTGGGAA 240
 QY 241 CCACCCGAGAGAGAGATTCCTTGTGAGAGAGATTCCTTGTGAGAGAGATTCCTT 300
 Db 241 CCACCCGAGAGAGAGATTCCTTGTGAGAGAGATTCCTTGTGAGAGAGATTCCTT 300
 QY 301 GAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGATCTACTGTGT 360
 Db 301 GAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGATCTACTGTGT 360
 QY 361 GAGGCTCTTGAAGATCTTCAAGAACCCGAGATTAATGAGAGAGAGAGAGAT 420
 Db 361 GAGGCTCTTGAAGATCTTCAAGAACCCGAGATTAATGAGAGAGAGAGAGAT 420
 QY 421 GACCAAGTCACTTGGCGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 Db 421 GACCAAGTCACTTGGCGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 481 GCGGCGCGCTTCAAGTCCCGGTGAGATTCGCGCCAGAGCTGCGCGCTTCTGCGCGCC 540
 Db 481 GCGGCGCGCTTCAAGTCCCGGTGAGATTCGCGCCAGAGCTGCGCGCTTCTGCGCGCC 540
 QY 541 CTGCGCGCGCTTGAAGCTCTGAGGCTTCAAGCTCCCGCGCTTCCGAGAACTGCGCGCC 600
 Db 541 CTGCGCGCGCTTGAAGCTCTGAGGCTTCAAGCTCCCGCGCTTCCGAGAACTGCGCGCC 600
 QY 601 AACCAATGGCCACAGTGTGCACTGACCTGCTCTGAGGCTAGAGAGAGAGAGAG 660
 Db 601 AACCAATGGCCACAGTGTGCACTGACCTGCTCTGAGGCTAGAGAGAGAGAGAG 660
 QY 661 GGGCGGAGATCCGCGCTTGTGAGTGTGCACTGAGAGAGAGAGAGAGAGAGAGAG 720
 Db 661 GGGCGGAGATCCGCGCTTGTGAGTGTGCACTGAGAGAGAGAGAGAGAGAGAGAG 720
 QY 721 TCGGAGCACACTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 Db 721 TCGGAGCACACTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 QY 781 ACCGCTTTGCAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 Db 781 ACCGCTTTGCAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 QY 841 GCGTTTCTGAG 900
 Db 841 GCGTTTCTGAG 900

Db 841 GCGTTTCTGAG 900
 QY 901 GAGAAATCGGTGAG 960
 Db 901 GAGAAATCGGTGAG 960
 QY 961 CTGCGCTCTGAG 1020
 Db 961 CTGCGCTCTGAG 1020
 QY 1021 GCGGAG 1080
 Db 1021 GCGGAG 1080
 QY 1081 CAGACCCCTCTGAG 1140
 Db 1081 CAGACCCCTCTGAG 1140
 QY 1141 GCGAG 1200
 Db 1141 GCGAG 1200
 QY 1201 AGTCTGAG 1260
 Db 1201 AGTCTGAG 1260
 QY 1261 GCGCTGAG 1320
 Db 1261 GCGCTGAG 1320
 QY 1321 AGGAG 1380
 Db 1321 AGGAG 1380
 QY 1381 ACTGAG 1440
 Db 1381 ACTGAG 1440
 QY 1441 GAG 1500
 Db 1441 GAG 1500
 QY 1501 TTTTAAATTAATTTATTAAT 1522
 Db 1501 TTTTAAATTAATTTATTAAT 1522

RESULT 2
 US-09-967-237-1
 ; Sequence 1, Application US/09967237
 ; Publication No. US20030049828A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zavada, Jan
 ; APPLICANT: Pastorekova, Silvia
 ; TITLE OF INVENTION: MN Gene and Protein
 ; FILE REFERENCE: D-0021, SB-2
 ; CURRENT APPLICATION NUMBER: US/09/967,237
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: 09/178,115
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1522
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (13)..(1389)
 ; NAME/KEY: mat_peptide
 ; LOCATION: (124)..(1389)
 US-09-967-237-1

Query Match 100.0%; Score 1522; DB 10; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAGAGCAGCCGATGAGCTCCCTGTCGCCCCAGCCCTGGCTCCCTGCTGTTGATCCCGCC 60
DB 1 AAGAGCAGCCGATGAGCTCCCTGTCGCCCCAGCCCTGGCTCCCTGCTGTTGATCCCGCC 60
QY CCTGCTCAAGGCTCACTGTGCACTGTGCTGCTCACTGTGCTGCTGCTGCTGCTGCTGCT 120
DB CCTGCTCAAGGCTCACTGTGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 CCCCAGAGGTTGCCCCGATGTCAGAGGATTTCCCTTTGGAGAGAGGCTCTTTGGGGAA 180
DB 121 CCCCAGAGGTTGCCCCGATGTCAGAGGATTTCCCTTTGGAGAGAGGCTCTTTGGGGAA 180
QY 181 GATGACCCATGGGCGAGAGGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
DB 181 GATGACCCATGGGCGAGAGGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
QY 241 CCACCCGAGAGAGAGATCTTACCTGGAAGAGAGATCTTACCTGGAAGAGAGATCTTACCT 300
DB 241 CCACCCGAGAGAGAGATCTTACCTGGAAGAGAGATCTTACCTGGAAGAGAGATCTTACCT 300
QY 301 GAAGTTAAGCCTAAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGGATCTTACCTAGT 360
DB 301 GAAGTTAAGCCTAAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGGATCTTACCTAGT 360
QY 361 GAGGCTCTGAGATCTTCAAGAACCCAGAAATATGCCCAAGAGGACAAAGAGAGGAGAT 420
DB 361 GAGGCTCTGAGATCTTCAAGAACCCAGAAATATGCCCAAGAGGACAAAGAGAGGAGAT 420
QY 421 GACCCAGAGATCTGGGCGTATGAGAGCGACCCGCTGGCCCGGGGCTGCCCGAGCTGC 480
DB 421 GACCCAGAGATCTGGGCGTATGAGAGCGACCCGCTGGCCCGGGGCTGCCCGAGCTGC 480
QY 481 GCGGCGCGCTTCAGTCCCGGCTGATATCCGCCCCCACTGCGCGCTTCTGCGCGGCC 540
DB 481 GCGGCGCGCTTCAGTCCCGGCTGATATCCGCCCCCACTGCGCGCTTCTGCGCGGCC 540
QY 541 CTGCGCCCCCTGGAATCTTGGGCTTCCAGCTCCCGCGCTCCCAAGAACTGGCGCTCGCG 600
DB 541 CTGCGCCCCCTGGAATCTTGGGCTTCCAGCTCCCGCGCTCCCAAGAACTGGCGCTCGCG 600
QY 601 AACATGCGCACAGTGTGCACTGACCTGCTCTGCGGCTAGAGATGCTCTGCGTCCC 660
DB 601 AACATGCGCACAGTGTGCACTGACCTGCTCTGCGGCTAGAGATGCTCTGCGTCCC 660
QY 661 GGGCGGAGTACCGGCTCTGCACTGCTGCACTGCGGGGCTGCAAGTCTGCTCGCGGC 720
DB 661 GGGCGGAGTACCGGCTCTGCACTGCTGCACTGCGGGGCTGCAAGTCTGCTCGCGGC 720
QY 721 TCGGAGCACAGTGTGAGAGGACCGGTTCCCTGCGAGATTCAGAGTGTTCACCTAGC 780
DB 721 TCGGAGCACAGTGTGAGAGGACCGGTTCCCTGCGAGATTCAGAGTGTTCACCTAGC 780
QY 781 ACCGCTTTGCCAGATTCAGAGGCTTTGGGGCGCCCGGAGGCTTGCGCTGTGCTGCTG 840
DB 781 ACCGCTTTGCCAGATTCAGAGGCTTTGGGGCGCCCGGAGGCTTGCGCTGTGCTGCTG 840
QY 841 GCGCTTCTGAGAGAGGCGCGGAGAAACAGTGTGCTGAGAGGCTGCTGCTGCTGCTG 900
DB 841 GCGCTTCTGAGAGAGGCGCGGAGAAACAGTGTGCTGAGAGGCTGCTGCTGCTGCTGCTG 900
QY 901 GAAGAAATGCTGAGAGAGGCTCAGAGATCAGGCTCCAGAGATTCGAGATTCGAGATC 960
DB 901 GAAGAAATGCTGAGAGAGGCTCAGAGATCAGGCTCCAGAGATTCGAGATTCGAGATC 960
QY 961 CTGCGCTCTGACTTCAAGCGCTTCAATTAAGAGGCTCTGACTTCAACCGCCCTGT 1020
DB 961 CTGCGCTCTGACTTCAAGCGCTTCAATTAAGAGGCTCTGACTTCAACCGCCCTGT 1020

```

```

QY 1021 GCCCAGGCTGCTGATCTGCTGTTTAAACCAAGACAGTATGCTGAGTGAAGCAGCTC 1080
DB 1021 GCCCAGGCTGCTGATCTGCTGTTTAAACCAAGACAGTATGCTGAGTGAAGCAGCTC 1080
QY 1081 CACACCCCTCTGACACCCCTGTGGGAGCTGTGACTCTGAGCTGCTGAGCTGAGCTTCCGA 1140
DB 1081 CACACCCCTCTGACACCCCTGTGGGAGCTGTGACTCTGAGCTGCTGAGCTGAGCTTCCGA 1140
QY 1141 GCGACGAGCCTTTGAATGGGCGAGTATGAGGCTCTTCCCTGCTGAGTGAACAGC 1200
DB 1141 GCGACGAGCCTTTGAATGGGCGAGTATGAGGCTCTTCCCTGCTGAGTGAACAGC 1200
QY 1201 AGTCTGCGGCTGTGAGAGGAGTGAAGTGAAGGCTCTTCCCTGCTGAGTGAACAGC 1260
DB 1201 AGTCTGCGGCTGTGAGAGGAGTGAAGTGAAGGCTCTTCCCTGCTGAGTGAACAGC 1260
QY 1261 GCGCTGCTTTTGGCTCTTCTTCTGTCAGCAGCGCTGCTTCTTCTGTCAGATGAG 1320
DB 1261 GCGCTGCTTTTGGCTCTTCTTCTGTCAGCAGCGCTGCTTCTTCTGTCAGATGAG 1320
QY 1321 AGGAGGCAAGAGAGGAGACCAAGGGGGTGTAGCTTACCGCCAGAGAGTATGAGCGAG 1380
DB 1321 AGGAGGCAAGAGAGGAGACCAAGGGGGTGTAGCTTACCGCCAGAGAGTATGAGCGAG 1380
QY 1381 ACTGAGGCTGAGGCTGATCTTGAAGAAATGAGAGGAGGAGGAGGAGGATCTGAGG 1440
DB 1381 ACTGAGGCTGAGGCTGATCTTGAAGAAATGAGAGGAGGAGGAGGAGGATCTGAGG 1440
QY 1441 GAGGCGGTATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 1441 GAGGCGGTATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1501 TTTTAAATTAATTTTAAAT 1522
DB 1501 TTTTAAATTAATTTTAAAT 1522

RESULT 3
US-09-954-456-89
; Sequence 89, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954, 456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233, 617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234, 052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234, 923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235, 134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235, 637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235, 638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235, 711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 89
; LENGTH: 1552
; TYPE: DNA

```

ORGANISM: Homo sapiens
US-09-954-456-89

Query Match 100.0%; Score 1522; DB 9; Length 1552;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ACAGTCAGCCGAGTGGCTCCCTGTGCCCCAGCCCTGCTCCCTGTGATCCCGGCC 60
DB 31 ACAGTCAGCCGAGTGGCTCCCTGTGCCCCAGCCCTGCTCCCTGTGATCCCGGCC 90
QY 61 CCGTCCAGGCTCAGTGTGCACTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 120
DB 91 CCGTCCAGGCTCAGTGTGCACTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 150
QY 121 CCCGAGAGTGGCCCGGATGCAAGAGATTCCTTGTGAGAGAGGCTCTTCTGGGAA 180
DB 151 CCCGAGAGTGGCCCGGATGCAAGAGATTCCTTGTGAGAGAGGCTCTTCTGGGAA 210
QY 181 GATGACCCAGTGGGCGAGAGAGATTCCTGCAAGAGATTCACCCAGAGAGAGAT 240
DB 211 GATGACCCAGTGGGCGAGAGAGATTCCTGCAAGAGATTCACCCAGAGAGAGAT 270
QY 241 CCACCCGAGAGAGAGATCTACTGAGAGAGATCTACTGAGAGAGATCTACT 300
DB 271 CCACCCGAGAGAGAGATCTACTGAGAGAGATCTACTGAGAGAGATCTACT 330
QY 301 GAAGTTAAGCTTAATCAGAAAGAGGCTCCCTGAAGTTAAGAGATCTACTACT 360
DB 331 GAAGTTAAGCTTAATCAGAAAGAGGCTCCCTGAAGTTAAGAGATCTACTACT 390
QY 361 GAGGCTCTGAGAGATCTCAAGAAAGCCCAAGATATCCCAAGAGAGAGAGAGAT 420
DB 391 GAGGCTCTGAGAGATCTCAAGAAAGCCCAAGATATCCCAAGAGAGAGAGAGAT 450
QY 421 GACGAGATCAATGGCGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
DB 451 GACGAGATCAATGGCGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 510
QY 481 GGGGGCGGCTTCAGATCCCGGAGATATCCCGGAGATATCCCGGAGATATCC 540
DB 511 GGGGGCGGCTTCAGATCCCGGAGATATCCCGGAGATATCCCGGAGATATCC 570
QY 541 CTGGCCCCCTGGAATCTCTGAGCTTCAGCTCCCGGCTCCCAAGATCTGCTG 600
DB 571 CTGGCCCCCTGGAATCTCTGAGCTTCAGCTCCCGGCTCCCAAGATCTGCTG 630
QY 601 AACATGGCCAGAGTGCACACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 631 AACATGGCCAGAGTGCACACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690
QY 661 GGGCGGAGATCCGGGCTCTGAGTGCATGCACTGGGGGCTGAGGTCGTCGGGC 720
DB 691 GGGCGGAGATCCGGGCTCTGAGTGCATGCACTGGGGGCTGAGGTCGTCGGGC 750
QY 721 TGGAGAGACATGTGGAAGGCAACGCTTCCCTGCGAGATCCACGAGTTCACCT 780
DB 751 TGGAGAGACATGTGGAAGGCAACGCTTCCCTGCGAGATCCACGAGTTCACCT 810
QY 781 ACCGCTTTGCGAGAGTTCAGAGAGCTTGGGGGCGCCGGAGGCTGAGCTGTG 840
DB 811 ACCGCTTTGCGAGAGTTCAGAGAGCTTGGGGGCGCCGGAGGCTGAGCTGTG 870
QY 841 GCTTTTCTGAGAGAGGCGCCGGAAGAAAAGTGCCTATAGAGAGTTCGCTGCT 900
DB 871 GCTTTTCTGAGAGAGGCGCCGGAAGAAAAGTGCCTATAGAGAGTTCGCTGCT 930
QY 901 GAAGAATGCTGAGAGAGGCTCAGAGATCAGAGTCCAGAGATCAGATATCTGAC 960
DB 931 GAAGAATGCTGAGAGAGGCTCAGAGATCAGAGTCCAGAGATCAGATATCTGAC 990
QY 961 CTGGCTTCTGACTTCAGCGGCTACTTCCAAATATAGAGGCTTCTGATTAAC 1020
```

```
DB 991 CTGGCTTCTGACTTCAGCGGCTACTTCCAAATATAGAGGCTTCTGATTAAC 1050
QY 1021 GCCAGAGGCTGATCTGAGCTGTGTTTAAACAGAGATGATGATGCTAAGAGCTC 1080
DB 1051 GCCAGAGGCTGATCTGAGCTGTGTTTAAACAGAGATGATGATGCTAAGAGCTC 1110
QY 1081 CAGACCTCTCTGACACCTGTGGGAGACCTGTGTGCTGTGCTGTGCTGTGCTG 1140
DB 1111 CAGACCTCTCTGACACCTGTGGGAGACCTGTGTGCTGTGCTGTGCTGTGCTG 1170
QY 1141 GCGACGAGCTTTGATGAGGCGAGTGAATTGAGGCTCTTCCCTGCTGAGAGAG 1200
DB 1171 GCGACGAGCTTTGATGAGGCGAGTGAATTGAGGCTCTTCCCTGCTGAGAGAG 1230
QY 1201 AGTCTCTGGGCTGTGAGCAAGTCCAGCTGAATTCCTGCTGAGCTGTGAGAT 1260
DB 1231 AGTCTCTGGGCTGTGAGCAAGTCCAGCTGAATTCCTGCTGAGCTGTGAGAT 1290
QY 1261 GCCCTGATTTTGGCTCTCTTTTGTGCTGACAGAGCTGCGCTTCTGTGAGAGA 1320
DB 1291 GCCCTGATTTTGGCTCTCTTTTGTGCTGACAGAGCTGCGCTTCTGTGAGAGA 1350
QY 1321 AGGAGACAGAAAGGGAACCAAGGGGCTGTGAGCTACCGCCAGAGAGTACCG 1380
DB 1351 AGGAGACAGAAAGGGAACCAAGGGGCTGTGAGCTACCGCCAGAGAGTACCG 1410
QY 1381 ACTGAGACCTAGAGGCTGTGATCTTGGAGAAATGAGAAAGCCAGCAGAGCAT 1440
DB 1411 ACTGAGACCTAGAGGCTGTGATCTTGGAGAAATGAGAAAGCCAGCAGAGCAT 1470
QY 1441 GAGCGGCTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 1471 GAGCGGCTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1530
QY 1501 TTTTAAATTAATTAATTAAT 1522
DB 1531 TTTTAAATTAATTAATTAAT 1552
```

RESULT 4
US-09-954-456-726
Sequence 726, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 726

```

; LENGTH: 1552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-726

```

Query Match	100.0%	Score 1522	DB 9	Length 1552
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1522; Conservative	0	Mismatches	0	Indels 0; Gaps 0

Qy	ACAGTCAGACGCAATGAGCTCCCTGTGGCCCAAGCCCTGAGCTCCTCTGTGATCCCGAGCC	60
Db	ACAGTCAGACGCAATGAGCTCCCTGTGGCCCAAGCCCTGAGCTCCTCTGTGATCCCGAGCC	90
Qy	CTGTCTCCAGGCTCACTGTGCACCTGCTGTGTCACTGCTGCTTCTGATCCTGTCCAT	120
Db	CTGTCTCCAGGCTCACTGTGCACCTGCTGTGTCACTGCTGCTTCTGATCCTGTCCAT	150
Qy	CCCCAGAGGTTGCCCCCGAGTGTGAGAGGATTCGCCCTTGGAGAGAGGCTTTCTGGAGAA	180
Db	CCCCAGAGGTTGCCCCCGAGTGTGAGAGGATTCGCCCTTGGAGAGAGGCTTTCTGGAGAA	210
Qy	GATGACCCACTGGGCGAGAGGATCTGCTCCAGTAAAGAGATTCACCCAGAGAGAGAT	240
Db	GATGACCCACTGGGCGAGAGGATCTGCTCCAGTAAAGAGATTCACCCAGAGAGAGAT	270
Qy	CCACCCGGAGAGAGGATTACTACTGGAGAAGAGATCTACTGGAGAGAGATCTACTCT	300
Db	CCACCCGGAGAGAGGATTACTACTGGAGAAGAGATCTACTGGAGAGAGATCTACTCT	330
Qy	GAAGTTAAGCTTAATCAGAGAGAGGAGCTCCCTGAAGTTAGAGATCTACTACTGT	360
Db	GAAGTTAAGCTTAATCAGAGAGAGGAGCTCCCTGAAGTTAGAGATCTACTACTGT	390
Qy	GAGGCTCTCTGAGATTCCTTAAGAACCCCAAGATTAATGCCCCAGAGGCAAGAAGGAGAT	420
Db	GAGGCTCTCTGAGATTCCTTAAGAACCCCAAGATTAATGCCCCAGAGGCAAGAAGGAGAT	450
Qy	GACACAGATCATTTGGGCGCATGAGAGGACACCCGCTTGACCCCGAGGTCCCAAGCTGC	480
Db	GACACAGATCATTTGGGCGCATGAGAGGACACCCGCTTGACCCCGAGGTCCCAAGCTGC	510
Qy	GCGGCGCGCTTCCAGTCCCGGTGATATCCGCCCAAGCTGTGCGGCTTCTGTCCCGGCGC	540
Db	GCGGCGCGCTTCCAGTCCCGGTGATATCCGCCCAAGCTGTGCGGCTTCTGTCCCGGCGC	570
Qy	CTGCGCCCCCTGGAAATCCTGTGGGCTTCCAGCTCCCGCGCTCCAGAACTGCAGCTCGGC	600
Db	CTGCGCCCCCTGGAAATCCTGTGGGCTTCCAGCTCCCGCGCTCCAGAACTGCAGCTCGGC	630
Qy	AACAAATGCGCACAGTGTGCACTGACCCCTGCTCTGTGGCTAGAGATGTCTGTGGGTCCC	660
Db	AACAAATGCGCACAGTGTGCACTGACCCCTGCTCTGTGGCTAGAGATGTCTGTGGGTCCC	690
Qy	GGGCGGAGATACCGGAGCTCTGCGAGCTGCATCTGCACTGGGGGCGTGCAGGTCGTCCGGGC	720
Db	GGGCGGAGATACCGGAGCTCTGCGAGCTGCATCTGCACTGGGGGCGTGCAGGTCGTCCGGGC	750
Qy	TCGGAGACACATGTGGAAGGCAACCGTTCCTCTGCGAGATTCACAGTGTTCACCTCAGC	780
Db	TCGGAGACACATGTGGAAGGCAACCGTTCCTCTGCGAGATTCACAGTGTTCACCTCAGC	810
Qy	ACGCGCTTTGCGAGATTTGACAGAGGCTTTGGGGGCGCCCGGAGAGGCTGGCGTGTGGGCGC	840
Db	ACGCGCTTTGCGAGATTTGACAGAGGCTTTGGGGGCGCCCGGAGAGGCTGGCGTGTGGGCGC	870
Qy	GCTTTTCTGGAGAGGGGCTCCGGAAGAAAACAAGTGTATGAGCAATTGCTGTGCTCGTTG	900
Db	GCTTTTCTGGAGAGGGGCTCCGGAAGAAAACAAGTGTATGAGCAATTGCTGTGCTCGTTG	930
Qy	GAAGAAATTCGCTAGAGAGGCTCAGAGACTGACGTCCAGAGACTGAGCATATCTGCACTC	960
Db	GAAGAAATTCGCTAGAGAGGCTCAGAGACTGACGTCCAGAGACTGAGCATATCTGCACTC	990

QY	961	CTGCCCCGTGACTTCAGCCGCTACTCCAAATATGAGAGGGCTCTCTAGCAACACCCGCCGT	1020
Db	991	CTGCCCTCTGACTTCAGCCGCTACTTCCTCAATATGAGAGGGCTCTCTAGCAACACCCGCCGT	1050
QY	1021	GCCCAAGGGGTGTCATCTGTGACTGTGTTTAAACAGACAGTAGTCTGAGTGTCTAAGCAGCTC	1080
Db	1051	GCCCAAGGGGTGTCATCTGTGACTGTGTTTAAACAGACAGTAGTCTGAGTGTCTAAGCAGCTC	1110
QY	1081	CACACCCCTCTTGACACCCCTGTGGGACCTGGTGTGATCTTCGGCTACACTGAACTTCGGA	1140
Db	1111	CACACCCCTCTTGACACCCCTGTGGGACCTGGTGTGATCTTCGGCTACACTGAACTTCGGA	1170
QY	1141	GGAGCGACCCCTTGAAATGAGCGAGTAGTATGAGAGCCCTCTTCCCTGCTGAGTGTGACAGC	1200
Db	1171	GGAGCGACCCCTTGAAATGAGCGAGTAGTATGAGAGCCCTCTTCCCTGCTGAGTGTGACAGC	1230
QY	1201	AGTCTCTGAGGCTGTGAGCCAGTCTCAGCTGAATTCCTGCTGAGCTGTGTGACATCTTA	1260
Db	1231	AGTCTCTGAGGCTGTGAGCCAGTCTCAGCTGAATTCCTGCTGAGCTGTGTGACATCTTA	1290
QY	1261	GCCCTGTGTTTGTGCTCTCTTTTGTGCTGTACACAGGCTGGCTTCTTGTGTGAGATGGA	1320
Db	1291	GCCCTGTGTTTGTGCTCTCTTTTGTGCTGTACACAGGCTGGCTTCTTGTGTGAGATGGA	1350
QY	1321	AGGACGACAGAGGGGAGACCAAGAGGGGCTGTGAGCTTACCGCCACAGAGGTATGCCAG	1380
Db	1351	AGGACGACAGAGGGGAGACCAAGAGGGGCTGTGAGCTTACCGCCACAGAGGTATGCCAG	1410
QY	1381	ACTGAGGCTGTAAGGCTTGATCTTGGAGAAATGTGAGAACTCAGACCCAGACAGGCAATCTGAGG	1440
Db	1411	ACTGAGGCTGTAAGGCTTGATCTTGGAGAAATGTGAGAACTCAGACCCAGACAGGCAATCTGAGG	1470
QY	1441	GAGAGCGGTAACTGTCTCTGTCTGTCTATTAATGCCACTTCTTAACGTGCCAAGAAAT	1500
Db	1471	GAGAGCGGTAACTGTCTCTGTCTGTCTATTAATGCCACTTCTTAACGTGCCAAGAAAT	1530
QY	1501	TTTTAAATTAATATTTAAT 1522	
Db	1531	TTTTAAATTAATATTTAAT 1552	

```

RESULT 5
US-09-960-706-1080
; Sequence 1080, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Mungert, William B.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyper
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1080
; LENGTH: 1552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 X66839
US-09-960-706-1080

```

	Query Match	100.0%	Score 1522	DB 10	Length 1552
	Best Local Similarity	100.0%	Pred. No. 0		
	Matches 1522	Conservative 0	Mismatches 0	Indels 0	Gaps 0
OY	1	A C A G T C A G C G C G A T G C T C C C C T T G T G C C C A G C C C C T G A C T C C C T T G T A T A T C C C G A G C C	60		
Dd	31	A C A G T A G C C G A T G C T C C C C T T G T G C C C A G C C C C T G A C T T G T A T A T C C C G A C C	90		

121 CCCAGAGGTTGCCCCGATGCAAGAGATTTCCCTTGGAGAGAGGCTTTCTGAGGAA 180
151 CCCAGAGGTTGCCCCGATGCAAGAGATTTCCCTTGGAGAGAGGCTTTCTGAGGAA 210
181 GATACACCACTGGGCGGAGAGAGATTTCTGAGAGAGAGATTTCCCTTGGAGAGAGAT 240
211 GATACACCACTGGGCGGAGAGAGATTTCTGAGAGAGAGATTTCCCTTGGAGAGAGAT 270
241 CCACCCGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 300
271 CCACCCGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 330
301 GAAGTTAAGCTTAATCAGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAT 360
331 GAAGTTAAGCTTAATCAGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAT 390
361 GAGGCTCTGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAT 420
391 GAGGCTCTGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAT 450
421 GACCAAGATCAATGGGCGGATGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 480
451 GACCAAGATCAATGGGCGGATGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 510
481 GGGGCGGCTTCCAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 540
511 GGGGCGGCTTCCAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 570
541 CTGCGCCCTTGGAACTCTGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGATTT 600
571 CTGCGCCCTTGGAACTCTGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGATTT 630
601 AACATGACCAAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 660
631 AACATGACCAAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 690
661 GGGGCGGATTTCCAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 720
691 GGGGCGGATTTCCAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 750
721 TCGAGACACTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAG 780
751 TCGAGACACTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAG 810
781 ACCGCTTTGACAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 840
811 ACCGCTTTGACAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 870
841 GCTTTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 900
871 GCTTTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 930
901 GAAAGAAATCGATGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAG 960
931 GAAAGAAATCGATGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAG 990
961 CTGCGCTTCTGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 1020
991 CTGCGCTTCTGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 1050
1021 GCGCAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGAT 1080
1051 GCGCAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGAT 1110
1081 CACACCTTCTGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 1140
1111 CACACCTTCTGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAG 1170
1141 GCGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGAT 1200
1171 GCGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGAT 1230
1201 AGTCTCGGAGATTTCTGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGAT 1260

1231 AGTCTCGGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAG 1290
1261 GCGCAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGATTTCTGAG 1320
1291 GCGCAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGATTTCTGAG 1350
1321 AGGACAGACAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 1380
1351 AGGACAGACAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAG 1410
1381 ACTGAGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAG 1440
1411 ACTGAGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAG 1470
1441 GAGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTT 1500
1471 GAGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTTCTGAGAGAGAGATTT 1530
1501 TTTTAAATTAATTTATTTAT 1522
1531 TTTTAAATTAATTTATTTAT 1552

RESULT 7
US-10-301-822-11
Sequence 11, Application US/10301822
Publication No. US20030148410A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Berger, Allison
APPLICANT: Guillemette, Tracy L.
APPLICANT: Kamatkar, Shubhang
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Thibodeau, Stephen N.
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF COLON CANCER
FILE REFERENCE: PPM01-029P2RNM
CURRENT FILING DATE: US/10/301,822
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/339,971
PRIOR FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: US 60/361,978
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/381,988
PRIOR FILING DATE: 2002-05-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 1552
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (43)...(1422)
US-10-301-822-11

Query Match 100.0%; Score 1522; DB 14; Length 1552;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ACAGTCCGAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAG 60
31 ACAGTCCGAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAG 90
61 CCGTCCAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAT 120
91 CCGTCCAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAT 150
121 CCCAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAGATTTCTGAGAGAT 180

D	b	151	CCCCAAGGTTCCCCCGAATGCAGAGAGATTC	CCCCCTTGGAGAGAGGCTCTTGGGGAA	210			
Q	y	181	GATGACCCACTGGGCGAGAGAGATCTGC	CCAGTGAAGAGATTCA	CCAGAGAGAGAT	240		
D	b	211	GATGACCCACTGGGCGAGAGAGATCTGCC	AGTGAAGAGATTCA	CCAGAGAGAGAT	270		
Q	y	241	CCACCCGAGAGAGAGATCTACCTG	AGAGAGATCTACCTG	AGAGAGAGAT	300		
D	b	271	CCACCCGAGAGAGAGATCTACCTG	AGAGAGATCTACCTG	AGAGAGAGAT	330		
Q	y	301	GAAGTTAAGCTTAATCAGAAAGAGAGG	CTCCCTTAAGTTAAGATTTAC	TA	360		
D	b	331	GAAGTTAAGCTTAATCAGAAAGAGAGG	CTCCCTTAAGTTAAGATTTAC	TA	390		
Q	y	361	GAGGCTCTGAGAGATCTCAAGAACCC	CAGATTAATCCACAGGAGCAAA	AGAGGAT	420		
D	b	391	GAGGCTCTGAGAGATCTCTCAAGAACCC	CAGATTAATCCACAGGAGCAAA	AGAGGAT	450		
Q	y	421	GACCAAGATCATTTGGCGCTATG	AGAGGCA	CCCCGCTTG	CCCCAGCTGC	480	
D	b	451	GACCAAGATCATTTGGCGCTATG	AGAGGCA	CCCCGCTTG	CCCCAGCTGC	510	
Q	y	481	GCGGGCGGCTTCAGTCCCCG	TGATTCG	CCCGCAGCTCG	CCGCTTCG	540	
D	b	511	GCGGGCGGCTTCAGTCCCCG	TGATTCG	CCCGCAGCTCG	CCGCTTCG	570	
Q	y	541	CTGGCGCCCTGGA	CTCTTG	GGCTTCAGGCT	CCCCGCTTC	600	
D	b	571	CTGGCGCCCTGGA	CTCTTG	GGCTTCAGGCT	CCCCGCTTC	630	
Q	y	601	AACAAATGGCCACAGTGTGCA	CTGACCTG	CTG	CGCTAGAGATG	CGCTCGGATCC	660
D	b	631	AACAAATGGCCACAGTGTGCA	CTGACCTG	CTG	CGCTAGAGATG	CGCTCGGATCC	690
Q	y	661	GGGCGGAGATACCGG	CTTGTG	CACTGATCTG	CACTGGGGGCTG	CAAGTGTCCGGG	720
D	b	691	GGGCGGAGATACCGG	CTTGTG	CACTGATCTG	CACTGGGGGCTG	CAAGTGTCCGGG	750
Q	y	721	TCGGAGACA	CTGTG	AGAGGAC	CGATTCCTCTG	CCGAGATTC	780
D	b	751	TCGGAGACA	CTGTG	AGAGGAC	CGATTCCTCTG	CCGAGATTC	810
Q	y	781	ACCGCTTTGCCAGAGTTA	CGAGGCTTGG	GGGCGGCGGAGGCTTGG	CGCTGTTGGC	840	
D	b	811	ACCGCTTTGCCAGAGTTA	CGAGGCTTGG	GGGCGGCGGAGGCTTGG	CGCTGTTGGC	870	
Q	y	841	GCCCTTGTGAGAGAGG	CCCGAGAA	AACAGTCC	ATAGAGATG	CTGCTCGCTTG	900
D	b	871	GCCCTTGTGAGAGAGG	CCCGAGAA	AACAGTCC	ATAGAGATG	CTGCTCGCTTG	930
Q	y	901	GAGAAATCGCTTA	GAGAAAGCTCA	AGACTCA	GGTCCCAGAGATG	GCATATCTG	960
D	b	931	GAGAAATCGCTTA	GAGAAAGCTCA	AGACTCA	GGTCCCAGAGATG	GCATATCTG	990
Q	y	961	CTGCGCTCTACTTAC	CGCGCTACTTCC	ATATAGAGG	AGTCTCTG	CTACACCGGCTGT	1020
D	b	991	CTGCGCTCTACTTAC	CGCGCTACTTCC	ATATAGAGG	AGTCTCTG	CTACACCGGCTGT	1050
Q	y	1021	GCCCAAGGCTCATCTG	GA	CTGTGTTAAC	CAGACAGTGA	TGCTGA	1080
D	b	1051	GCCCAAGGCTCATCTG	GA	CTGTGTTAAC	CAGACAGTGA	TGCTGA	1110
Q	y	1081	CACACCTCTCTAC	CACCCCTTG	GGGAGCTCG	GTGACTCTCG	GCCTACAGCTGA	1140
D	b	1111	CACACCTCTCTAC	CACCCCTTG	GGGAGCTCG	GTGACTCTCG	GCCTACAGCTGA	1170
Q	y	1141	GCGAGCAGAGCTTTGA	ATGGG	CGAGTGA	GCCTCTCCCTG	CTGGAGTGA	1200
D	b	1171	GCGAGCAGAGCTTTGA	ATGGG	CGAGTGA	GCCTCTCCCTG	CTGGAGTGA	1230
Q	y	1201	AGTCTG	GGGCTG	TAGGCA	GTGCA	GTGCTG	1260

Db	1231	AGTCTCGGCGTGTGAGCCAGTCCACACTGAATCTCGCTGAGCTGGTGAATCCTTA	1290
Qy	1261	GGCCTGATTTTGGCTTCCTTTTGTGCTGTACCAAGCGTGGCTTCTGTGAGATGAGA	1320
Db	1291	GGCCTGGATTTTGGCTTCCTTTTGTGCTGTACCAAGCGTGGCTTCTGTGAGATGAGA	1350
Qy	1321	AGGACGACAGAGGGGAGCCAAAGGGGGTGTAGCTACCGCCACAGAGGTAGCCGAG	1380
Db	1351	AGGACGACAGAGGGGAGCCAAAGGGGGTGTAGCTACCGCCACAGAGGTAGCCGAG	1410
Qy	1381	ACTGAGGCTTAAGGCTTGAATCTTGAGAAATGGAABAACGACGACAGAGGATCTGAGG	1440
Db	1411	ACTGAGGCTTAAGGCTTGAATCTTGAGAAATGGAABAACGACGACAGAGGATCTGAGG	1470
Qy	1441	GGAGCGGTAATGTCTGTCTGCTCATTAATGCACTTCCTTTAACTGCCAGAAAT	1500
Db	1471	GGAGCGGTAATGTCTGTCTGCTCATTAATGCACTTCCTTTAACTGCCAGAAAT	1530
Qy	1501	TTTAAATAAATAATTTAAT 1522	
Db	1531	TTTAAATAAATAATTTAAT 1552	

RESULT 8

US-10-465-572-9

Sequence 9, Application US/10465572

Publication No.: US20030207840A1

GENERAL INFORMATION:

APPLICANT: Riggins, Gregory

APPLICANT: Lal, Anita

TITLE OF INVENTION: GENES INDUCED BY HYPOXIA

FILE REFERENCE: 000250.00012

CURRENT APPLICATION NUMBER: US/10/465.572

PRIOR FILING DATE: 2003-06-20

PRIOR APPLICATION NUMBER: US/10/201.642

PRIOR FILING DATE: 2002-07-24/6/307,600

PRIOR APPLICATION NUMBER: 60/307,600

PRIOR FILING DATE: 2001-07-26

NUMBER OF SEQ ID NOS: 30.

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9

LENGTH: 1552

TYPE: DNA

ORGANISM: Homo sapiens

US-10-465-572-9

Query Match 100.0%; Score 1522; DB 15; Length 1552;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ACAGTCAGCGCGATGAGTCCCTGTCGCCCCAGCCCTGAGCTCCCTGCTGATCCCGAGCC	60
Db	31	ACAGTCAGCGCGATGAGTCCCTGTCGCCCCAGCCCTGAGCTCCCTGCTGATCCCGAGCC	90
Qy	61	CTGTCTCAGGCGCTCACTGTGTCACTGTCTGTCACTGTCTTGTGATGCTGTTCAT	120
Db	91	CTGTCTCAGGCGCTCACTGTGTCACTGTCTGTCACTGTCTTGTGATGCTGTTCAT	150
Qy	121	CCCCAGAGGTTGCCCGAGATGTGACGAGAGATTCCTCCCTTGGAGAGAGGCTTTCTGGAGAA	180
Db	151	CCCCAGAGGTTGCCCGAGATGTGACGAGAGATTCCTCCCTTGGAGAGAGGCTTTCTGGAGAA	210
Qy	181	GATGACCCACTGGGCGAGAGAGATTTGCCACGTGAAGAAGATTCACCCAGAGAGAGAT	240
Db	211	GATGACCCACTGGGCGAGAGAGATTTGCCACGTGAAGAAGATTCACCCAGAGAGAGAT	270
Qy	241	CCACCCGAGAGAGAGATTTACTTGAGAGAGAGATCTACCTGGAGAGAGATCTAACC	300
Db	271	CCACCCGAGAGAGAGATTTACTTGAGAGAGAGATCTAACC	330
Qy	301	GAGTTTAAGCTTAATCAAGAAAGAGAGGCTCCCTGAAGTTAGAGATCTAAGTCTAAGT	360
Db	331	GAGTTTAAGCTTAATCAAGAAAGAGAGGCTCCCTGAAGTTAGAGATCTAAGTCTAAGT	390

QY 361 GAGGCTCTGAGATCTCAAGAACCCAGATAATATGCCCACAGAGCAAAAGAGGGAT 420
DB 391 GAGGCTCTGAGATCTCTCAAGAACCCAGATAATATGCCCACAGAGCAAAAGAGGGAT 450
QY 421 GACCAAGATCATTTGGCGCTATGAGAGCAACCGCCCTGGCCCCGGGTGTCCCAAGCTGC 480
DB 451 GACCAAGATCATTTGGCGCTATGAGAGCAACCGCCCTGGCCCCGGGTGTCCCAAGCTGC 510
QY 481 GCGGGCCGCTGCAATCCCGGGTGAATATCCGCCCCAGCTGCGCGCTTCTGGCCGCGC 540
DB 511 GCGGGCCGCTTCAAGTCCCGGGTGAATATCCGCCCCAGCTGCGCGCTTCTGGCCGCGC 570
QY 541 CTGCGCCCCCTGGAATCTCGGGCTTCAAGCTCCCGCGCTTCCAGAGACTGCGCTCGC 630
DB 571 CTGCGCCCCCTGGAATCTCGGGCTTCAAGCTCCCGCGCTTCCAGAGACTGCGCTCGC 630
QY 601 AACATATGCGACAGTGTGCACTGACCCCGCTTCTGGCTAGAGATGCTCTGGGCTCC 660
DB 631 AACATATGCGACAGTGTGCACTGACCCCGCTTCTGGCTAGAGATGCTCTGGGCTCC 650
QY 661 GCGCGGAGTACCGGGCTCTGAGTGAATCTGCACTGAGGCGGCTGAGTCTGCGGCGC 720
DB 691 GCGCGGAGTACCGGGCTCTGAGTGAATCTGCACTGAGGCGGCTGAGTCTGCGGCGC 750
QY 721 TCGGAGCACACTGTGAAAGGCCACCGTTTCCCTGCGGAGATCCAGCGTGTCACTGAGC 780
DB 751 TCGGAGCACACTGTGAAAGGCCACCGTTTCCCTGCGGAGATCCAGCGTGTCACTGAGC 810
QY 781 ACCGCTTTGCCAGATTTGACAGAGGCTTGGGGCGCCCGGAGGCGCTGCGGCTGTTGGCC 840
DB 811 ACCGCTTTGCCAGATTTGACAGAGGCTTGGGGCGCCCGGAGGCGCTGCGGCTGTTGGCC 870
QY 841 GCGCTTTCTGAGAGAGGCGCCGGAAGAAACAATGCTATGAGCAATGCTCTGCGCTG 900
DB 871 GCGCTTTCTGAGAGAGGCGCCGGAAGAAACAATGCTATGAGCAATGCTCTGCGCTG 930
QY 901 GAAGAATCGCTGAGAGAGGCTCAGAGACTCAGAGTCCAGAGACTGAGATATCTGCACTC 960
DB 931 GAAGAATCGCTGAGAGAGGCTCAGAGACTCAGAGTCCAGAGACTGAGATATCTGCACTC 990
QY 961 CTGCGCTCTGACTTCAAGCGGCTATCTCAATATGAGGGGCTCTTCACTACACCGCCCTGT 1020
DB 991 CTGCGCTCTGACTTCAAGCGGCTATCTCAATATGAGGGGCTCTTCACTACACCGCCCTGT 1050
QY 1021 GCCCAGGGTGTCACTGAGACTGTGTTAAACAGACAGTATGCTGAGTGAAGAGCTC 1080
DB 1051 GCCCAGGGTGTCACTGAGACTGTGTTAAACAGACAGTATGCTGAGTGAAGAGCTC 1110
QY 1081 CACACCTCTCTGACACCCCTGTGGGAGCTGTGATCTGCGCTACAGCTGAATCTTCCGA 1140
DB 1111 CACACCTCTCTGACACCCCTGTGGGAGCTGTGATCTGCGCTACAGCTGAATCTTCCGA 1170
QY 1141 GCGACGAGGCTTGAATGGGAGGAGTGAAGGCTCTTCCCTGCTGAGTGAAGCAGC 1200
DB 1171 GCGACGAGGCTTGAATGGGAGGAGTGAAGGCTCTTCCCTGCTGAGTGAAGCAGC 1230
QY 1201 AGTCTCGGGCTCTGAGCCAGTCAAGCTGAATCTGCTGCTGCTGCTGCTGCTGCTA 1260
DB 1231 AGTCTCGGGCTCTGAGCCAGTCAAGCTGAATCTGCTGCTGCTGCTGCTGCTGCTA 1290
QY 1261 GCGCTGCTTTTGGCCCTCTTTTGTGTGACCAAGAGTGTGCTTCTTGTGAGATGAGA 1320
DB 1291 GCGCTGCTTTTGGCCCTCTTTTGTGTGACCAAGAGTGTGCTTCTTGTGAGATGAGA 1350
QY 1321 AGGAGACACAGAAAGGGAGCAAAAGGGGTGTGAGTACCGCCACAGCAGAGTACCGAG 1380
DB 1351 AGGAGACACAGAAAGGGAGCAAAAGGGGTGTGAGTACCGCCACAGCAGAGTACCGAG 1410
QY 1381 ACTGAGACCTTAGAGGCTGTGATCTTGTGAGATGTGAGAAAGCCAGAGGAGTCTGAGAGG 1440
DB 1411 ACTGAGACCTTAGAGGCTGTGATCTTGTGAGATGTGAGAAAGCCAGAGGAGTCTGAGAGG 1470

QY 1441 GAGCGCGGTAACTGTCTGCTGCTGCTCATATGCACTTCTTTAAGTCCAGAAATT 1500
DB 1471 GAGCGCGGTAACTGTCTGCTGCTGCTCATATGCACTTCTTTAAGTCCAGAAATT 1530
QY 1501 TTTTAAATTAATTTAAT 1522
DB 1531 TTTTAAATTAATTTAAT 1552

RESULT 9
US-10-295-027-305
Sequence 305, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afari, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Gineberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: US/10/295,027
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,345
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 305
LENGTH: 1552
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-305

Query Match 100.0%; Score 1522; DB 15; Length 1552;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGTCAGCGGATGAGTCCCTGTGCCCCAGCCCTGCTCCCTGTGTATCCGAGCC 60
DB 31 ACAGTCAGCGGATGAGTCCCTGTGCCCCAGCCCTGCTCCCTGTGTATCCGAGCC 90
QY 61 CTGTCTCAGGCTCTCACTGTGTGAACTGTCTGTCTGCTGCTGCTTTGAAAGCCGTCCAT 120
DB 91 CTGTCTCAGGCTCTCACTGTGTGAACTGTCTGTCTGCTGCTGCTTTGAAAGCCGTCCAT 150
QY 121 CCCACAGGTTGCCCGGAGTGCAGAGAGATTCCCTTTGGAGAGGCTCTTCTGGGGA 180
DB 151 CCCACAGGTTGCCCGGAGTGCAGAGAGATTCCCTTTGGAGAGGCTCTTCTGGGGA 210

QY 181 GATGACCACTGGGGGAGAGAGATCTGCCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
DB 211 GATGACCACTGGGGGAGAGAGATCTGCCCAGTGAAGAGATTCACCCAGAGAGAGAT 270
QY 241 CCACCCGGAGAGAGAGATCTGAGAGAGAGATCTACCTGGAGAGAGATCTACCT 300
DB 271 CCACCCGGAGAGAGAGATCTGAGAGAGAGATCTACCTGGAGAGAGATCTACCT 330
QY 301 GAAGTTAAGCTTAAATCAGAAAGAGAGAGAGATCTGAGAGATCTACCTGTT 360
DB 331 GAAGTTAAGCTTAAATCAGAAAGAGAGAGAGATCTGAGAGATCTACCTGTT 390
QY 361 GAGGCTCTGAGAGATCTCAGAAAGAGAGAGATCTGAGAGAGATCTGAGAGAGAT 420
DB 391 GAGGCTCTGAGAGATCTCAGAAAGAGAGAGATCTGAGAGAGATCTGAGAGAGAT 450
QY 421 GACCAAGATCATGAGAGATCTGAGAGAGAGATCTGAGAGAGATCTGAGAGATCTGC 480
DB 451 GACCAAGATCATGAGAGATCTGAGAGAGAGATCTGAGAGAGATCTGAGAGATCTGC 510
QY 481 GGGGGCCGCTTCCAGTCCCGGTGATATCCGCCCCAGCTCGCGGCTTCTGCGGGCC 540
DB 511 GGGGGCCGCTTCCAGTCCCGGTGATATCCGCCCCAGCTCGCGGCTTCTGCGGGCC 570
QY 541 CTGCGCCCTGGAATCCTGAGGCTTCCAGCTCCCGGCTCCAGAACTGCGCTGCG 600
DB 571 CTGCGCCCTGGAATCCTGAGGCTTCCAGCTCCCGGCTCCAGAACTGCGCTGCG 630
QY 601 AACATGCGCAAGTGTGCAATGACCTGCTGCTGAGGCTGAGAGATGAGTCTGAGTCCC 660
DB 631 AACATGCGCAAGTGTGCAATGACCTGCTGCTGAGGCTGAGAGATGAGTCTGAGTCCC 690
QY 661 GGGGGGAGATCCGGGCTGCTGAGGCTGATGCACTGAGGAGCTGAGGCTGAGGCTG 720
DB 691 GGGGGGAGATCCGGGCTGCTGAGGCTGATGCACTGAGGAGCTGAGGCTGAGGCTG 750
QY 721 TGGAGACACTGTGAG 780
DB 751 TGGAGACACTGTGAG 810
QY 781 ACCGCTTGGAGAGATGAG 840
DB 811 ACCGCTTGGAGAGATGAG 870
QY 841 GCGTTTCTGAG 900
DB 871 GCGTTTCTGAG 930
QY 901 GAAGAATGCTGAG 960
DB 931 GAAGAATGCTGAG 990
QY 961 CTGCGCTTGAATCAGCGGCTACTTCCAAATGAGGAGCTTCTGAGTCAACCGGCTGT 1020
DB 991 CTGCGCTTGAATCAGCGGCTACTTCCAAATGAGGAGCTTCTGAGTCAACCGGCTGT 1050
QY 1021 GCGCAGAGTGTATCTGAG 1080
DB 1051 GCGCAGAGTGTATCTGAG 1110
QY 1081 CACACCTCTCTGACACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1111 CACACCTCTCTGACACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170
QY 1141 GCGAG 1200
DB 1171 GCGAG 1230
QY 1201 AGTCTCTGAG 1260
DB 1231 AGTCTCTGAG 1290

QY 1261 GCGCTGTTTGGAG 1320
DB 1291 GCGCTGTTTGGAG 1350
QY 1321 AGGAG 1380
DB 1351 AGGAG 1410
QY 1381 ACTGAG 1440
DB 1411 ACTGAG 1470
QY 1441 GAG 1500
DB 1471 GAG 1530
QY 1501 TTTTAAATTAATTAATTAAT 1552
DB 1531 TTTTAAATTAATTAATTAAT 1552

RESULT 10

US-10-295-027-1022

Sequence 1022, Application US/10295027

Publication No. US20030232350A1

GENERAL INFORMATION:

APPLICANT: Afar, Daniel

APPLICANT: Aziz, Natasha

APPLICANT: Ginsberg, Wendy M.

APPLICANT: Glynn, Richard

APPLICANT: Heyez, Peter A.

APPLICANT: Mack, David H.

APPLICANT: Murray, Richard

APPLICANT: Watson, Susan R.

APPLICANT: Eos Biotechnology, Inc.

TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

FILE REFERENCE: 018501-012500US

CURRENT APPLICATION NUMBER: US/10/295,027

PRIORITY FILING DATE: 2002-11-13

PRIORITY FILING DATE: 2000-09-15

PRIORITY FILING DATE: 2000-09-15

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

PRIORITY FILING DATE: 2001-11-13

Query Match 100.0%; Score 1522; DB 15; Length 1552;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ACAGTACAGCGAGTACGGGTGCTCCCTGTGTGCCCCAGGCGCTGAGTCCCTCTGTGTGATCCCGAGC	60
Db	31	ACAGTACAGCGAGTACGGGTGCTCCCTGTGTGCCCCAGGCGCTGAGTCCCTCTGTGTGATCCCGAGC	90
Qy	61	CTGTGTCAGAGGCTCACTGTGTGCACTGTGTGTCACTGTGCTTTGTATGCTGTTCAT	120
Db	91	CTGTGTCAGAGGCTCACTGTGTGCACTGTGTGTCACTGTGTGCTTTGTATGCTGTTCAT	150
Qy	121	CCCCAGAGGTGTGCCCGGATGTGACAGAGGATTTCCGCCCTTGAGAGAGAGTCTTCTGTGGGAA	180
Db	151	CCCCAGAGGTGTGCCCGGATGTGACAGAGGATTTCCGCCCTTGAGAGAGAGTCTTCTGTGGGAA	210
Qy	181	GATGACCCCATGTGSCGAGAGAGGATCTGTCCCATGTGAAGAGATTCACTCCAGAGAGAGAT	240
Db	211	GATGACCCCATGTGSCGAGAGAGGATCTGTCCCATGTGAAGAGATTCACTCCAGAGAGAGAT	270
Qy	241	CCACCCGGAGAGAGGATCTTACCTGTGAGAGAGGATTTACCTGTGAGAGAGATTTACCT	300
Db	271	CCACCCGGAGAGAGGATCTTACCTGTGAGAGAGGATTTACCTGTGAGAGAGATTTACCT	330
Qy	301	GAGTTAAGCTTAATACAAAGAGAGGCTCCCTGAAGTTAAGAGATTTACCTACTGTT	360
Db	331	GAGTTAAGCTTAATACAAAGAGAGGCTCCCTGAAGTTAAGAGATTTACCTACTGTT	390
Qy	361	GAGGTCTGTGAGATTCCTCAAGAACCCAGATATATCCACAGAGGACAAAGAGAGGAT	420
Db	391	GAGGTCTGTGAGATTCCTCAAGAACCCAGATATATCCACAGAGGACAAAGAGAGGAT	450
Qy	421	GACCAAGTCAATTGGCGCTATGAGAGGGAACCCGCTCTGGCCCGGGGTGTCCCAAGCTGC	480
Db	451	GACCAAGTCAATTGGCGCTATGAGAGGGAACCCGCTCTGGCCCGGGGTGTCCCAAGCTGC	510
Qy	481	GCGGCGCGCTTCCAGTCCCGGAGATATCCGCGCCCAAGCTGCGCGCTTCTGTGCGCGC	540
Db	511	GCGGCGCGCTTCCAGTCCCGGAGATATCCGCGCCCAAGCTGCGCGCTTCTGTGCGCGC	570
Qy	541	CTGCGCGCGCTTGGAACTCTGTGGGCTTCCAGTCTCCCGCGCTCCCAAGACTGGCGCTGGCG	600
Db	571	CTGCGCGCGCTTGGAACTCTGTGGGCTTCCAGTCTCCCGCGCTCCCAAGACTGGCGCTGGCG	630
Qy	601	AACAATGCGCACAGTGTGCAACTGACCTGTGCTCTGTGGCTTGAAGATGGCTGTGGATCCC	660
Db	631	AACAATGCGCACAGTGTGCAACTGACCTGTGCTCTGTGGCTTGAAGATGGCTGTGGATCCC	690
Qy	661	GCGCGGAGATACCGGGCTCTGTGAGTGTGATCTGCATCTGGGGGCTTGAAGGTCTCCGGGC	720
Db	691	GCGCGGAGATACCGGGCTCTGTGAGTGTGATCTGCATCTGGGGGCTTGAAGGTCTCCGGGC	750
Qy	721	TGCGAGACACATGTGGAAGAGCCACCGTTCCTGTCCGAGATCAAGTGGTTCACTGAGC	780
Db	751	TGCGAGACACATGTGGAAGAGCCACCGTTCCTGTCCGAGATCAAGTGGTTCACTGAGC	810
Qy	781	ACCGCTTTGCCAGATGTGACGAGGCGCTTGGGGGCGCCGGAGGCGTGGCGGTGTGGGC	840
Db	811	ACCGCTTTGCCAGATGTGACGAGGCGCTTGGGGGCGCCGGAGGCGTGGCGGTGTGGGC	870
Qy	841	GCGTTTCTGAGAGAGGCGCCGAGAGAAACATGTGCTATGACAGATTGCTGTGCGTTG	900
Db	871	GCGTTTCTGAGAGAGGCGCCGAGAGAAACATGTGCTATGACAGATTGCTGTGCGTTG	930
Qy	901	GAAAGAAATGGCTGAGAGAGGCTCAGAGACTCAGGTCCAGAGACTGGAATATCTGCACCTC	960
Db	931	GAAAGAAATGGCTGAGAGAGGCTCAGAGACTCAGGTCCAGAGACTGGAATATCTGCACCTC	990
Qy	961	CTGCGCTGTGACTTCAACCGCGTACTTCCAAATGAGAGGCTCTTGACTTCAACCGCGCTGT	1020
Db	991	CTGCGCTGTGACTTCAACCGCGTACTTCCAAATGAGAGGCTCTTGACTTCAACCGCGCTGT	1050
Qy	1021	GCCGAGAGTGCATCTGAGACTGTGTTTAAACGACAGAGATCTGAGTGAAGCAGCTC	1080
Db	1051	GCCGAGAGTGCATCTGAGACTGTGTTTAAACGACAGAGATCTGAGTGAAGCAGCTC	1110

QY	1081	CACACCCCTCTGACACCCCTGTGGGAGCCGTGTGACTCTCGGCTACAGCTAACTTCGGA	1140
Db	1111	CACACCCCTCTGTACACCCCTGTGGGAGCCGTGTGACTCTCGGCTACAGCTAACTTCGGA	1170
QY	1141	GGGAGCGACCCCTTTGAATGGGCGAGATTTGAAGGCTCTCTCCCTGCTGTGAGTGGACAGC	1200
Db	1171	GGAGCGACCCCTTTGAATGGGCGAGATTTGAAGGCTCTCTCCCTGCTGTGAGTGGACAGC	1230
QY	1201	AGTCCTCGGGGCTGTGAGCCAGTGCACGTGAATTCCTGCTGACTGCTGTGACATCTTA	1260
Db	1231	AGTCCTCGGGGCTGTGAGCCAGTGCACGTGAATTCCTGCTGACTGCTGTGACATCTTA	1290
QY	1261	GGCCTGTGTTTTGTGGCCTCTTTTGTCTGTACACAGGCTGCGCTTCTGTGTGACATGAGA	1320
Db	1291	GGCCTGTGTTTTGTGGCCTCTTTTGTCTGTACACAGGCTGCGCTTCTGTGTGACATGAGA	1350
QY	1321	AGGCAAGCACAGAAAGGGGAAACCAAGGGGGGTGTAGCTAACCGCCACAGAGGTATACAG	1380
Db	1351	AGGCAAGCACAGAAAGGGGAAACCAAGGGGGGTGTAGCTAACCGCCACAGAGGTATACAG	1410
QY	1381	ACTGAGCCTTAGAGGCTGTGATCTTTGAGAAATGTGAGAGCCAGCCAGAGCATCTGAGGG	1440
Db	1411	ACTGAGCCTTAGAGGCTGTGATCTTTGAGAAATGTGAGAGCCAGCCAGAGCATCTGAGGG	1470
QY	1441	GGAGCCGGTAACTGTCTCTGTCTCTGCTCATTTATGCACTTCTCTTTAACTGGCAAGAAATT	1500
Db	1471	GGAGCCGGTAACTGTCTCTGTCTCTGCTCATTTATGCACTTCTCTTTAACTGGCAAGAAATT	1530
QY	1501	TTTTAAATTAATATTTTAAT	1522
Db	1531	TTTTAAATTAATATTTTAAT	1552

```

RESULT 11 RESULT
US-10-102-524-1695
Sequence 1695, Application US/10102524
Publication No. US2003010943A1
GENERAL INFORMATION:
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Galager, Alexander
APPLICANT: Gordon, Brian
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT OF KIDNEY CANCER
FILE REFERENCE: 210121.572
CURRENT APPLICATION NUMBER: US/10/102.524
CURRENT FILING DATE: 2002-03-19
NUMBER OF SEQ ID NOS: 1863
SOFTWARE: FASTSEQ FOR Windows Version 4.0
SEQ ID NO 1695
LENGTH: 1519
TYPE: DNA
ORGANISM: Homo sapiens
US-10-102-524-1695

```

Query Match	99.7%	Score 1517.4	DB 14	Length 1519
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1516	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

[illegible]

Db 799 CCTCAAGAACCCAGAAATATATGCCACAGGAGCAAGAGGAGATGACCAAGATCATTTGG 858
 Qy 436 CGCATATGAGGAGCCCGCCCTGAGCCCGGGGTGTCCCGAGCTTGCGCGGGCGCTTCCAG 495
 Db 859 CGCTATGAGGAGCCCGCCCTGAGCCCGGGGTGTCCCGAGCTTGCGCGGGCGCTTCCAG 918
 Qy 496 TCCCGGATGATATCCGCCCCCAGCTGCGGCTTGTGCCCCGAGCTTGCGCGGGCGCTTCCAG 555
 Db 919 TCCCGGATGATATCCGCCCCCAGCTGCGGCTTGTGCCCCGAGCTTGCGCGGGCGCTTCCAG 978
 Qy 556 CTCCGAGCTTCCAGCTGCGGCTTCCAGAGCTGCGGCTTCCAGAGCTGCGGCTTCCAGAGCT 615
 Db 979 CTCCGAGCTTCCAGCTGCGGCTTCCAGAGCTGCGGCTTCCAGAGCTGCGGCTTCCAGAGCT 1038
 Qy 616 GTGCACTGACCCCTGCTCTGAGGCTAGAGATGCTGAGGCTGCGGCGGGAGTACCGG 675
 Db 1039 GTGCACTGACCCCTGCTCTGAGGCTAGAGATGCTGAGGCTGCGGCGGGAGTACCGG 1098
 Qy 676 GCTTGTGAGCTGATCTGCACTGCGGGGCTGCGAGTGTCCGGGCTTCCAGAGCACTGTG 735
 Db 1099 GCTTGTGAGCTGATCTGCACTGCGGGGCTGCGAGTGTCCGGGCTTCCAGAGCACTGTG 1158
 Qy 736 GAGGCGACCGTTTCCCTGCGAGATCCAGTGTGCTTACCTCAAGACCGCTTCCAGAG 795
 Db 1159 GAGGCGACCGTTTCCCTGCGAGATCCAGTGTGCTTACCTCAAGACCGCTTCCAGAG 1218
 Qy 796 GTTACGAGGCTTGGGGGCGCGCGGAGGCTTGGCGGCTTGGCGGCTTGGAGAGAG 855
 Db 1219 GTTACGAGGCTTGGGGGCGCGCGGAGGCTTGGCGGCTTGGCGGCTTGGAGAGAG 1278
 Qy 856 GCGCCGAGAGAAACAGTGTGCTTACAGATGCTGCTGCTTGGAGAGAAATCCGTGAG 915
 Db 1279 GCGCCGAGAGAAACAGTGTGCTTACAGATGCTGCTGCTTGGAGAGAAATCCGTGAG 1338
 Qy 916 GAAGGCTCAGAGACTCAGGTCCTCCAGAGCTGGAATATCTGCACTCTGCGCTTGTACTTC 975
 Db 1339 GAAGGCTCAGAGACTCAGGTCCTCCAGAGCTGGAATATCTGCACTCTGCGCTTGTACTTC 1398
 Qy 976 AGCGGCTACTTCCAAATGAGAGGCTCTGACTTACACCGGCTGTGCCAGGGGTGCATC 1035
 Db 1399 AGCGGCTACTTCCAAATGAGAGGCTCTGACTTACACCGGCTGTGCCAGGGGTGCATC 1458
 Qy 1036 TGAAGTGTATTAAACCAAGAGATGCTGATGCTGATGAGAGCTTCCAGACCTCTGTGAC 1095
 Db 1459 TGAAGTGTATTAAACCAAGAGATGCTGATGCTGATGAGAGCTTCCAGACCTCTGTGAC 1518
 Qy 1096 ACCCTGAGGAGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1155
 Db 1519 ACCCTGAGGAGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1578
 Qy 1156 AATGGGAGATGATGAGGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1215
 Db 1579 AATGGGAGATGATGAGGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1638
 Qy 1216 GAGGAGCTCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1275
 Db 1639 GAGGAGCTCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1698
 Qy 1276 CTCTTTTGTCTGCTCAGAGGCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1335
 Db 1699 CTCTTTTGTCTGCTCAGAGGCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1758
 Qy 1336 GGAACCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1395
 Db 1759 GGAACCAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1818
 Qy 1396 CTG 1398
 Db 1819 GTG 1821

RESULT 13
 US-09-772-719-5

; Sequence 5, Application US/09772719
 ; Patent No. US20020137910A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zavada, Jan
 ; APPLICANT: Pastorekova, Silvia
 ; APPLICANT: Pastorek, Jaromir
 ; TITLE OF INVENTION: MN Gene and Protein
 ; NUMBER OF SEQUENCES: 86
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Leona L. Lauder
 ; STREET: 369 Pine Street
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/772,719
 ; FILING DATE: 30-JAN-2001
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/485,049
 ; FILING DATE: 07-JUN-1995
 ; APPLICATION NUMBER: US 08/260,190
 ; FILING DATE: 15-JUN-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lauder, Leona L.
 ; REGISTRATION NUMBER: 30,863
 ; REFERENCE/DOCKET NUMBER: D-0021.3E
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-981-2034
 ; TELEFAX: 415-981-0332
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 10898 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULAR TYPE: DNA (genomic)
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; US-09-772-719-5
 Query Match 27.2%; Score 414.4; DB 9; Length 10898;
 Best Local Similarity 99.8%; Pred. No. 9.3e-112;
 Matches 415; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 ACHGTACGCGCATGAGTCTCCCTGTGTGCCCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 Db 3537 ACHGTACGCGCATGAGTCTCCCTGTGTGCCCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3596
 Qy 61 CCGCTCCAGGCTCTCACTGTGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 Db 3597 CCGCTCCAGGCTCTCACTGTGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3656
 Qy 121 CCCGAGAGTTGCCCCGAGATGCAAGAGATTTCCCTTGGAGAGAGGCTCTTCTGAGGAA 180
 Db 3657 CCCGAGAGTTGCCCCGAGATGCAAGAGATTTCCCTTGGAGAGAGGCTCTTCTGAGGAA 3716
 Qy 181 GATGACCACTGGGCGAGAGATCTGCCAGTGAAGAGATTCAACCGAGAGAGAGAT 240
 Db 3717 GATGACCACTGGGCGAGAGATCTGCCAGTGAAGAGATTCAACCGAGAGAGAGAT 3776
 Qy 241 CCACCGAGAGAGAGATTTACTCTGAGAGAGAGATTTACTCTGAGAGAGAGATTTACTCT 300
 Db 3777 CCACCGAGAGAGAGATTTACTCTGAGAGAGAGATTTACTCTGAGAGAGAGATTTACTCT 3836
 Qy 301 GAAGTTAGCTAAATCAGAGAGAGAGGCTCCCTGAAGTGAAGATTTACTCTGCTGTT 360

Mon Mar 8 10:16:30 2004

us-09-967-237-1.rnpb

Page 14

Db 383/ GAAGTTAAACCTAAATTCAGAAAGAGAGGGCCCTCGAAGTTAAGAGTCTACTACTGT 3896
 QY 361 GAGGCGCTCGAGATCTCCAGAACCCAGATAATGCCACAGGGAACAAGAGG 416
 Db 3897 GAGGCTCCCGAGATCTCAAGAACCCAGATAATGCCACAGGGAACAAGAGG 3952

RESULT 14
US-09-967-237-5

Sequence 5, Application US/09967237
Publication No. US20030043928A1
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
FILE REFERENCE: D-0021.5B-2
CURRENT APPLICATION NUMBER: US/09/967,237
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 09/1178,115
PRIOR FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 116
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 10898
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: gene
LOCATION: (1) ..(10898)
OTHER INFORMATION: full-length MN genomic sequence
NAME/KEY: unsure of base at position 1974
LOCATION: (1974)
OTHER INFORMATION: unsure of base at position 1974, which is in the 5' region flanking the start site of the MN gene.
OTHER INFORMATION: transcription initiation site (3507) as determined by RNase protection assay.
US-09-967-237-5

Query Match	27.2%	Score 414.4	DB 10	Length 10898
Best Local Similarity	99.8%	Pred. No. 9.3e-112		
Matches 415	Conservative	0	Mismatches 1	Indels 0
			Gaps 0	
QY	1	ACAGTCAGCCGATGATGCCCTCCCTGTGCCCCACCCCTGAGCTCCCTCTGTGATCCCGGCC	60	
Db	3537	ACAGTCAGCCGATGATGCCCTCCCTGTGCCCCACCCCTGAGCTCCCTCTGTGATCCCGGCC	3596	
QY	61	CTGTCTCCAGGCTCACTGTGCATGTCTGTTCATCTGCTCTTCTGATGCCCTGCAT	120	
Db	3597	CTGTCTCCAGGCTCACTGTGCATGTCTGTTCATCTGCTCTTCTGATGCCCTGCAT	3656	
QY	121	CCCCAGAGGTGTGCCCGGATGTGCAGAGAGATTCCCCCTGGAGAGAGCTCTTCTGGAGAA	180	
Db	3657	CCCCAGAGGTGTGCCCGGATGTGCAGAGAGATTCCCCCTGGAGAGAGCTCTTCTGGAGAA	3716	
QY	181	GATGACCCACTGTGGCGAGAGAGATCTGCCACGTGAAGAGGATTACCCAGAGAGAGAT	240	
Db	3717	GATGACCCACTGTGGCGAGAGAGATCTGCCACGTGAAGAGGATTACCCAGAGAGAGAT	3767	
QY	241	CCACCCGGAGAGAGAGATCTACTCTGGAAGAGAGATCTACTCTGAGAGAGATCTACT	300	
Db	3777	CCACCCGGAGAGAGAGATCTACTCTGGAAGAGAGATCTACTCTGAGAGAGATCTACT	3833	
QY	301	GAAGTTAAGCTTAATCAGAGAGAGAGGAGCTCCCTGAGATTAGAGATCTACCTACTGTT	360	
Db	3837	GAAGTTAAGCTTAATCAGAGAGAGAGGAGCTCCCTGAGATTAGAGATCTACTACTGTT	3896	
QY	361	GAGGCTCTGTGAAATCTCTAAGAACCCCAAGAAATATGCTCCACAGGAGCAAAAGAG	416	
Db	3897	GAGGCTCTGTGAAATCTCTAAGAACCCCAAGAAATATGCTCCACAGGAGCAAAAGAG	3952	

Patent No. US20020137910A1
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. lauder
STREET: 369 Pine Street
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/772,715
FILING DATE: 30-JAN-2001
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO.: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: 1st MN exon
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-772-719-28

Query	Match	Local Similarity	Score	DB	Length
27.2%	99.8%	Pred. No. 4,8e-112;			
Best Local Similarity	99.8%	Pred. No. 4,8e-112;			
Matches	414;	Conservative	0;	Mismatches	1;
				Indels	0;
				Gaps	0;

Db 301 GAAGTTAAGCCTAATCAGAAGAGAGGCTCCCTGAAGTTAGAGATCTACCTACTGTT 360
Qy 361 GAGGCTCTGAGATCTCTCAAGAAACCCAGATATATGCCCCACAGGACAAAGAG 415
Db 361 GAGGCTCTGAGATCTCTCAAGAAACCCAGATATATGCCCCACAGGACAAAGAG 415

Search completed: March 7, 2004, 06:33:57
Job time : 422 secs

http://fulllength.invitrogen.com/InvitrogenCorporation_1600
 Faraday Avenue Genoscope sequence ID : CS0D1085C09QPI.
 Location/Qualifiers

FEATURES

Source

1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1085F18"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 69.2%; Score 1053.8; DB 9; Length 1201;
 Best Local Similarity 96.2%; Pred. No. 3,7e-224;
 Matches 1086; Conservative 21; Mismatches 19; Indels 3; Gaps 3;

1 ACAGTCAGCCGATGCTCCCTGTCGCCAGCCCTGCTGCTCTGTTGATCCGAGC 60
 75 AAGATCAGCCGATGCTCCCTGTCGCCAGCCCTGCTGCTCTGTTGATCCGAGC 133
 61 CCTGCTCAGCCGATGCTCCCTGTCGCCAGCCCTGCTGCTCTGTTGATCCGAGC 120
 134 CCTGCTCAGCCGATGCTCCCTGTCGCCAGCCCTGCTGCTCTGTTGATCCGAGC 193
 121 CCCGAGAGGTGCTCCCGATGAGAGGATTCCTCCCTGAGAGAGGCTCTTGGGGAA 180
 194 CCCGAGAGGTGCTCCCGATGAGAGGATTCCTCCCTGAGAGAGGCTCTTGGGGAA 253
 181 GATGACCACTGGGAGAGAGATTCGCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
 254 GATGACCACTGGGAGAGAGATTCGCCAGTGAAGAGATTCACCCAGAGAGAGAT 313
 241 CCACCCGAGAGAGATTCCTGAGAGAGAGATTCCTGAGAGAGAGATTCCTGAG 300
 314 CCACCCGAGAGAGATTCCTGAGAGAGAGATTCCTGAGAGAGAGATTCCTGAG 373
 301 GAAGTTAAGCTTAATCAGAGAGAGAGATTCCTGAGAGAGATTCCTGAGAGAT 360
 374 GAAGTTAAGCTTAATCAGAGAGAGAGATTCCTGAGAGAGATTCCTGAGAGAT 433
 361 GAGGCTCTGAGAGATTCCTGAGAGAGAGATTCCTGAGAGAGAGATTCCTGAG 420
 434 GAGGCTCTGAGAGATTCCTGAGAGAGAGATTCCTGAGAGAGAGATTCCTGAG 493
 421 GACCAAGTCAATGGGCTATGAGAGAGAGATTCCTGAGAGAGATTCCTGAGAG 480
 494 GACCAAGTCAATGGGCTATGAGAGAGAGATTCCTGAGAGAGATTCCTGAGAG 553
 481 GCGGCGCGCTTCCAGTCCCGGCTGATTCGCGCCCGAGCTGCGGCTTTCGCG 540
 554 GCGGCGCGCTTCCAGTCCCGGCTGATTCGCGCCCGAGCTGCGGCTTTCGCG 613
 541 CTGCGCCCGCTGGAATCTCTGGGCTTCCAGTCCCGGCTGATTCGCGCCCGAG 600
 614 CTGCGCCCGCTGGAATCTCTGGGCTTCCAGTCCCGGCTGATTCGCGCCCGAG 673
 601 AACATGGCCAGAGTGTGCACTGACCTGCTCTGCTGCTGCTGCTGCTGCTGCT 660
 674 AACATGGCCAGAGTGTGCACTGACCTGCTCTGCTGCTGCTGCTGCTGCTGCT 733
 661 GGGCGGAGATACGGGCTCTGAGAGTGTGCACTGAGAGAGAGATTCCTGAGAG 720
 734 GGGCGGAGATACGGGCTCTGAGAGTGTGCACTGAGAGAGAGATTCCTGAGAG 793
 721 TCGAGAGCACTGTGAGAGAGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 794 TCGAGAGCACTGTGAGAGAGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 853

QY 781 ACCGCTTTGCCAGAGTTCAGAGGCTTTGGGAGCCCGGAGAGGCTGCTGTTGAGC 840

DB 854 ACCGCTTTGCCAGAGTTCAGAGGCTTTGGGAGCCCGGAGAGGCTGCTGTTGAGC 913

QY 841 GCTTTCTGAGAGAGGCGCCGAGAGAAACAGTGGCTATGAGCACTGCTGCTGCTG 900

DB 914 GCTTTCTGAGAGAGGCGCCGAGAGAAACAGTGGCTATGAGCACTGCTGCTGCTG 973

QY 901 GAGAAATCCGCTGAGAGAGGCTCAGAGTCCAGAGC-TGACATATCTGCACT 959

DB 974 GAGAAATCCGCTGAGAGAGGCTCAGAGTCCAGAGC-TGACATATCTGCACT 1033

QY 960 CTGTCCTCTGACTTACAGCCGCTACTTCCAAATATGAGGGTCTGACTACCGCCCTG 1019

DB 1034 CTGTCCTCTGACTTACAGCCGCTACTTCCAAATATGAGGGTCTGACTACCGCCCTG 1093

QY 1020 TGCCAGAGTGTATCTGAGTGTGTTTACAGAGAGTATGATGATGATGATGATGAT 1079

DB 1094 TGCCAGAGTGTATCTGAGTGTGTTTACAGAGAGTATGATGATGATGATGATGAT 1153

QY 1080 CCACACCTCTCTGACACCTCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1128

DB 1154 CCACACCTCTCTGACACCTCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1201

RESULT 2
 AL554705 1201 bp mRNA linear EST 31-MAY-2003

LOCUS AL554705 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

DEFINITION clone CS0D1085F18 5-PRIME, mRNA sequence.

ACCESSION AL554705

VERSION AL554705.2 GI:31276515

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Eukaryota; Euteleostomi; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1201)

AUTHORS Li W.B., Gruber C., Jesses J., and Polayes D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On Feb 15, 2001 this sequence version replaced GI:12895743.

Contact: Genoscope - Centre National de Sequencage

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 5300.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cg1-bin/cluster.cgi?seq=CS0D1085D09QPI&cluster=5300.f. Contact :

Feng Liang Email: fliang@lifetech.com URL :

http://fulllength.invitrogen.com/InvitrogenCorporation_1600

Faraday Avenue Genoscope sequence ID : CS0D1085C09QPI.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1085F18"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 60.5%; Score 920.8; DB 9; Length 1201;
 Best Local Similarity 96.4%; Pred. No. 1.8e-194;
 Matches 991; Conservative 13; Mismatches 15; Indels 9; Gaps 6;

1 ACAGTCAGCCGATGCTCCCTGTCGCCAGCCCTGCTGCTCTGTTGATCCGAGC 60


```

Db      |||
99  ACAGTAGCGGATGCTCCCTGTGCCCCAGCCCCCTGGCTCCCTCTGTATCCCGCC 158
Qy      |||
61  CCGCTCCAGGCTCTCTGTGAACGCTGCTGTACTGCTGCTCTCTGAAGCCGTCAT 120
Db      |||
159  CCGTCTCAGGCTCTCTGTGAACGCTGCTGTACTGCTGCTCTCTGTATCCCT 218
Qy      |||
121  CCCCAGAGGTTGCCCCGATGAGAGAGATTCCTCCCTTGGAGAGGCTCTTCTGGGAA 180
Db      |||
219  CCCCAGAGGTTGCCCCGATGAGAGAGATTCCTCCCTTGGAGAGGCTCTTCTGGGAA 278
Qy      |||
181  GATGACCCATCTGGGCGAGAGATCTGCCCACTGAGAGAGATTCACCCAGAGAGAT 240
Db      |||
279  GATGACCCATCTGGGCGAGAGATCTGCCCACTGAGAGAGATTCACCCAGAGAGAT 338
Qy      |||
241  CCACCCGAGAGAGATCTACTGAGAGAGAGATCTACTGAGAGAGAGATCTACT 300
Db      |||
339  CCACCCGAGAGAGATCTACTGAGAGAGAGATCTACTGAGAGAGAGATCTACT 398
Qy      |||
301  GAAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAGATTGAGAGATCTACTACT 360
Db      |||
399  GAAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAGATTGAGAGATCTACTACT 458
Qy      |||
361  GAGGCTCTGAGAGATCTCAAGAACCCAGAAATATGCCCCAGAGAGAGAGAGAT 420
Db      |||
459  GAGGCTCTGAGAGATCTCAAGAACCCAGAAATATGCCCCAGAGAGAGAGAGAT 518
Qy      |||
421  GACCAGAGTATTGGCGCTATGAGAGGCGACCCGCTGGCCCGCGTGTCTCCAGCT 480
Db      |||
519  GACCAGAGTATTGGCGCTATGAGAGGCGACCCGCTGGCCCGCGTGTCTCCAGCT 578
Qy      |||
481  GCGGCGCGCTTCCAGTCCCGCGGATATCCCGCCAGCTGCGCGCTTCCGCGGCG 540
Db      |||
579  GCGGCGCGCTTCCAGTCCCGCGGATATCCCGCCAGCTGCGCGCTTCCGCGGCG 638
Qy      |||
541  CTGGGCCCCCTGGAATCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGGCGCT 600
Db      |||
639  CTGGGCCCCCTGGAATCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGGCGCT 698
Qy      |||
601  AACATGCGCAAGTGTGCAATGACCTGCTGCTGCTGCTGAGATGCTGCTGGCTCC 660
Db      |||
699  AACATGCGCAAGTGTGCAATGACCTGCTGCTGCTGCTGAGATGCTGCTGGCTCC 758
Qy      |||
661  GCGGCGGAGTACCGGCTCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
Db      |||
759  GCGGCGGAGTACCGGCTCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
Qy      |||
720  CTGGAGACATCTGTGAGAGGCGACCGCTTCCCTGCTGCTGCTGCTGCTGCTGCT 779
Db      |||
819  CTGGAGACATCTGTGAGAGGCGACCGCTTCCCTGCTGCTGCTGCTGCTGCTGCT 878
Qy      |||
780  CACGCGCTTCCAGAGTGTGACAGAGGCTTGGGCGCGCCGAGAGGCTGCTGCTGCT 839
Db      |||
879  CACGCGCTTCCAGAGTGTGACAGAGGCTTGGGCGCGCCGAGAGGCTGCTGCTGCT 938
Qy      |||
840  CGCTTTCTGTGAGAGAGGCGCGAGAGAAACAGTGCCTATGAGCACTGCTGCTGCT 899
Db      |||
939  CGCTTTCTGTGAGAGAGGCGCGAGAGAAACAGTGCCTATGAGCACTGCTGCTGCT 997
Qy      |||
900  GGAAGAAATGCTGAGAGAGGCTCAGAGATCCAGAGTCCAGAGTCCAGAGTCCAGAT 959
Db      |||
998  GGAAGAAATGCTGAGAGAGGCTCAGAGATCCAGAGTCCAGAGTCCAGAGTCCAGAT 1056
Qy      |||
960  CTTGCGCTCTGACTTACCGCGCTACTTCCAAATATGAGGAGTCTCTGATCAACCGCT 1019
Db      |||
1057  C-TGGCTCTGACTTACG-CGCTAATTCAMAMWAGAGGCTCT---KATWMACCGCGCT 1110
Qy      |||
1020  TGGCCAGG 1027
Db      |||
1111  TGGCAGGG 1118

```

```

BX423970
LOCUS      BX423970      1013 bp      mRNA      linear      EST 15-MAY-2003
DEFINITION BX423970 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            CS0DA003YB12 5-PRIME, mRNA sequence.
ACCESSION  BX423970
VERSION     BX423970.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 1013)
AUTHORS   Li W.B., Gruber C., Jesse J., and Polayes D.
TITLES    Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT    Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 Evry cedex - France
            Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 5300.f. For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CSIDA001ZF09QPl&cluster=5300.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CSIDA001ZF09QPl.

FEATURES
    source
        1..1013
            location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DA003YB12"
                /tissue_type="NEUROBLASTOMA"
                /clone_id="Homo sapiens NEUROBLASTOMA"
                /note=Vector: pCMVSPORT 6; 1st strand cDNA was primed
                with a NotI-oligo(dT) primer. Five prime end enriched,
                double-strand cDNA was digested with Not I and cloned into
                the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                Library was not normalized.

ORIGIN
Query Match      60.1%; Score 915; DB 13; Length 1013;
Best Local Similarity 94.8%; Pred. No. 3.2e-193;
Matches 906; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

2  CAGTCAGCGCATGCTCCCTGTGCCCCAGCCCCCTGCTCCCTCTGTGATCCCGGCC 61
Db      |||
56  CCGGAGTCCGATGCTCCCTGTGCCCCAGCCCCCTGCTCCCTCTGTGATCCCGGCC 115
Qy      |||
62  CTGCTCAAGGCTCACTGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Db      |||
116  CTGCTCAAGGCTCACTGTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175
Qy      |||
122  CCCAGAGTGGCCCCGATGCAAGAGATTCCTCCCTTGGAGAGAGCTCTTCTGGGAG 181
Db      |||
176  CCCAGAGTGGCCCCGATGCAAGAGATTCCTCCCTTGGAGAGAGCTCTTCTGGGAG 235
Qy      |||
182  ATGACCATGCGCGAGAGAGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGATC 241
Db      |||
236  ATGACCATGCGCGAGAGAGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGATC 295
Qy      |||
242  CACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATCTACCT 301
Db      |||
296  CACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATCTACCT 355
Qy      |||
302  AAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAGATTGAGAGATCTACTACTGTT 361
Db      |||
356  AAGTTAAGCTTAATCAGAGAGAGAGGCTCCCTGAGATTGAGAGATCTACTACTGTT 415
Qy      |||
362  AGGCTCTGAGAGATCCGAGAGAGAGAGAGATCCCAAGAGAGAGAGAGAGATG 421
Db      |||
416  AGGCTCTGAGAGATCCGAGAGAGAGAGAGATCCCAAGAGAGAGAGAGAGATG 475

```

```

QY 422 ACCAGAGTATTGGCGCTATGAGAGGACCCGCTGCGCCCGGGTGTCCCAAGCTTGG 461
Db 476 ACCAGAGTATTGGCGCGATGAGAGGACCCGCTGCGCCCGGGTGTACCAACAGCTGGC 535
QY 482 CGGGCGGCTTCCAGATCCCGGATGATATCCGCGCCAGCTCGCGCTTGTGCCGCGCC 541
Db 536 CGGGCGGCTTCCAGATCCCGGATGATATCCGCGCCAGCTCGCGCTTGTGCCGCGCC 595
QY 542 TGGCGCCCTGGAATCTCTGGGCTTCCAGCTCCGCGCTCCAGAACTGGCGCTGCGCA 601
Db 596 TGGCGCCCTGGAATCTCTGGGCTTCCAGCTCCGCGCTCCAGAACTGGCGCTGCGCM 655
QY 602 ACAATGGCCACAGTGTGCAACTGACCCCTGCTCCGCGGATGAGATGGCTTGGATCCCG 661
Db 656 ACAATGGCCACAGTGTGCAACTGACCCCTGCTCCGCGGATGAGATGGCTTGGATCCCG 715
QY 662 GGGCGGAGTACCGGGGCTTGGACAGCTGCACTGGGCGGCTGGCAGATGTCGGGCT 721
Db 716 GGGCGGAGTACCGGGGCTTGGACAGCTGCACTGGGCGGCTGGCAGATGTCGGGCT 775
QY 722 CGAGGACACTGTGAGAGGACCAAGCTTCCCTGCGAGATCCAGTGGTTCAGCTCAGCA 781
Db 776 CGAGGACACTGTGAGAGGACCAAGCTTCCCTGCGAGATCCAGTGGTTCAGCTCAGCA 835
QY 782 CCGCCTTTGGCAGAGTTGACGAGGCTTGGGCGGCGCGGAGGCTTGGCGCTTGGCG 841
Db 836 CCGCCTTTGGCAGAGTTGAGAGGCTTGGGCGGCGCGGAGGCTTGGCGCTTGGCG 895
QY 842 CTTTCTGAGAGAGGCGCGGAGAAAGAAAGATGCTTATGAGAGTGTCTGTCTGGCTGG 901
Db 896 CTTTCTGAGAGAGGCGCGGAGAAAGAAAGATGCTTATGAGAGTGTCTGTCTGGCTGG 955
QY 902 AAGAAATCGCTGAGAGGCTCAGAGCTCAGATCCCAAGCTGAGCATATCTGCACT 959
Db 956 RAAAAAMGCTGAGAGGCTCAGAGATTTAGATTTAGATTTAGATTTAGATTTAGATTT 1013

```

```

RESULT 4
AL558378 1072 bp mRNA linear EST 31-MAY-2003
LOCUS AL558378 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ005YK14 5-PRIME, mRNA sequence.
ACCESSION AL558378 GI:31280177
VERSION AL558378.2
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1072)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12902842.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5300.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ005BF07QPLcluster=5300.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DJ005BF07QPL.

```

FEATURES

```

Source
1..1072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ005YK14"

```

```

ORIGIN
/cell_line="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/clone_id="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

Query Match 55.4%; Score 903.4; DB 9; Length 1072;
Best Local Similarity 99.1%; Pred. No. 1,3e-190;
Matches 956; Conservative 4; Mismatches 0; Indels 5; Gaps 5;
1 ACAGTACGCGCATGAGCTCCCTGTGCGCCAGCCCTGAGCTCCCTGTGATCCCGGC 60
Db 94 ACAGTACGCGCATGAGCTCCCTGTGCGCCAGCCCTGAGCTCCCTGTGATCCCGGC 153
QY 61 CTTGCTCAGAGCTTCACTGTGCAACTGTCTGTCTGCTGCTTGTGATGCTGTCCAT 120
Db 154 CTTGCTCAGAGCTTCACTGTGCAACTGTCTGTCTGCTGCTTGTGATGCTGTCCAT 213
QY 121 CCCGAGAGTTGCCCGGATGACAGAGATTTCCCTTGGAGAGAGCTTCTGGAGAA 180
Db 214 CCCGAGAGTTGCCCGGATGACAGAGATTTCCCTTGGAGAGAGCTTCTGGAGAA 273
QY 181 GATGACCCACTGGGCGAGAGAGATCTCCAGTGAAGAGATTCACCCAGAGAGAGAT 240
Db 274 GATGACCCACTGGGCGAGAGATCTCCAGTGAAGAGATTCACCCAGAGAGAGAT 333
QY 241 CCACCCGAGAGAGAGATCTTACCTGAGAGAGAGATTAAGTGAAGAGATTAAGT 300
Db 334 CCACCCGAGAGAGAGATCTTACCTGAGAGAGAGATTAAGTGAAGAGATTAAGT 393
QY 301 GAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGATTTACTACT 360
Db 394 GAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGATTTACTACT 453
QY 361 GAGGCTCTGAGATTCCTCAAGAAACCCAGAAATATGCCAGAGGAGCAAGAGAGAT 420
Db 454 GAGGCTCTGAGATTCCTCAAGAAACCCAGAAATATGCCAGAGGAGCAAGAGAGAT 513
QY 421 GACACAGATCATTTGGCGCTATGAGAGGACCCGCGCTGCGCGGAGTGTCCCAAGCTGC 480
Db 514 GACACAGATCATTTGGCGCTATGAGAGGACCCGCGCTGCGCGGAGTGTCCCAAGCTGC 573
QY 481 GCGGCGCGCTTCAAGTCCCGGAGATATCCGCGCGAGCTCGCGCGCTTCTGCGGCG 540
Db 574 GCGGCGCGCTTCAAGTCCCGGAGATATCCGCGCGAGCTCGCGCGCTTCTGCGGCG 633
QY 541 CTGCGCGCGCTTGAATCTCTGAGCTCCGCGCTCCAGAACTGCGGCTGCG 600
Db 634 CTGCGCGCGCTTGAATCTCTGAGCTCCGCGCTCCAGAACTGCGGCTGCG 693
QY 601 AACATGGCCACAGTGTGCAACTGACCTGCTCTTGGCTTGAAGATGGCTTGGGCTCC 660
Db 694 AACATGGCCACAGTGTGCAACTGACCTGCTCTTGGCTTGAAGATGGCTTGGGCTCC 753
QY 661 GGGCGGAGTACCGGGGCTCTGAGGCTGATGTGACTGGGAGGCTGAGAGTGTGCGGCG 720
Db 754 GGGCGGAGTACCGGGGCTCTGAGGCTGATGTGACTGGGAGGCTGAGAGTGTGCGGCG 813
QY 721 TGGAGACACTGTGGAAGGCCACCGTTTCCCTGCCAGATTCACGTTTCACTGAC 780
Db 814 TGGAGACACTGTGGAAGGCCACCGTTTCCCTGCCAGATTCACGTTTCACTGAC 873
QY 781 ACCGCTTTGCGAGAGTTGACGAGGCTTGGGAGCGCCGAGAGGCTTGGCTGTTGGC 840
Db 874 ACCGCTTTGCGAGAGTTGACGAGGCTTGGGAGCGCCGAGAGGCTTGGCTGTTGGC 922
QY 841 GCGCTTTCTGAGAGAGGCGCCGAGAGAAACAGTGTCTATGACAGTGTCTGTCTGCTTG 900
Db 933 GCGCTTTCTGAGAGAGGCGCCGAGAG-AAACAGTGTCTATGACAGTGTCTGTCTGCTTG 991

```

QY	900	GAA	AATGCTGAG	GAAGCTC	CAGACACTAGGTCC	CAGAATCGA	CAATCTG	ACTC	960
	:	:	:	:	:	:	:	:	
Dd	992	GAA	-RAATGGCTGA	GAA-GCTCAG	AGCTCAG	KCCCCAG	ACTGGA	CAMAT-TGC	1048
	:	:	:	:	:	:	:	:	
QY	961	CTG	CCC	965					
	:	:	:	:	:	:	:	:	
Dd	1049	CTG	CC	1053					

RESULT	5	
LOCUS	AL554665/c	
DEFINITION	Al554665 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSDD1085YF18-3-PRIME, mRNA sequence.	1201 bp linear EST 31-MAY-2003
ACCESSION	AL554665	
VERSION	AL554665	
KEYWORDS	EST.	
SOURCE	AL554665.2 GI:31276475	
	Homo sapiens (human)	

```

FEATURES
Source
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CSOD1.045Y.78"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match
Best local similarity 92.0%; Pred. No.4,6e-182;
Matches 1027; Conservative 23; Mismatches 47; Indels 19; Gaps 13

QY 407 ACAAAGAAAGGAGATGACCAAGCATGTGGCGGTATGAGAGGACCCCGCCCTGAGCCCGG 466
DB 1108 AGAAAAAARGGGATACCAATATTTTCGCTATGAAAGG-C-ACCGCCCTTATGCCCG 105
QY 467 TGTCCCGCAGCTGTGCGCGCGCGCGCTTCACAGTCCCGGATGCGATATCCGCCCGCAGCTGCGG 526
DB 1050 TGTCCCGCAGCTGTGCG-C-GGCGCGCTHMACTCCCGCGAT-GATATCGCGCCCGCG-TYCGCG 994
QY 527 CCTTTCGCGCGCGCGCGCGCGCGCGCTGAACTCCTGAGGCTTCAGCTCCCGCGCTCCAC 586
DB 993 CCTTTCGCGCGG-CCTGCGCGCCCTTGA--CTCTGAGSTTCGVSATCCCGCGCGCTCCAG 937
QY 567 AACTGCGCGCTTCGGAACATAGCGCAACAGTGTGCACTAACCTGCTCCTCGCGGCTAGAGA 646
DB 936 TACTGCG-CCTGCGGAACATAGCGCAAGTGTGCACTGACCTGCTCTCTGCGCTTAGAGA 878

```

QY	644	TGGCTCTGGAGTCCCGGACGGAGATACCGGGCTCTGAGCTGATCTGCACTGGGGGCTG	706
Db	877	TGGCTCTGGAGTCCCGGACGGAGATACCGGGCTCTGAGCTGATCTGCACTGGGGGCTG	818
QY	707	CAGGTCTGTCGGGCTCGGAGCACATGTGGAAAGCCACGGTTTCCTGCGAGMATCCACG	766
Db	817	CAGGTCTGTCGGGCTCGGAGCACATGTGGAAAGCCACGGTTTCCTGCGAGMATCCACG	758
QY	767	TGGTTCACCTAGACAGACGGCTTTGCCAGATTGACAGAGGCTTGGGGGCGCCCGGAGGCC	826
Db	757	TGGTTCACCTAGACAGACGGCTTTGCCAGATTGACAGAGGCTTGGGGGCGCCCGGAGGCC	695
QY	827	TGGCGGTGTGGCGGCTTTGTGGAGGAGGGCCCGGAGAAACAGTGGCTATGAGCAGT	886
Db	697	TGGCGGTGTGGCGGCTTTGTGGAGGAGGGCCCGGAGAAACAGTGGCTATGAGCAGT	638
QY	887	TGCTGTCTCGTTGGAGAAATGCTGTAAGGAAGGCTCAAGACTGACGTGCCAGACT--	944
Db	637	GGCTGTCTCGTTGGAGAAATGCTGTAAGGAAGGCTCAAGACTGACGTGCCAGACTCG	578
QY	945	GGACATATGTGACATCCC--GGCCCTGACCTTAAGCCG--TACCTTCAATATAGGGGCT	1002
Db	577	CGACATCTGTGACATCCC--GGCCCTCTACCTTAAGCCGCCACTTCCAAATATAGGGGCT	518
QY	1003	CTGACTACACCGCCCTGTGCCAGAGGTGTCATCTGACTGTGTGTTAACCAAGACAGTATG	1062
Db	517	CTGACTACACCGCCCTGTGCCAGAGGTGTCATCTGACTGTGTGTTAACCAAGACAGTATG	458
QY	1063	CTGAGTGTAAACACACTCCCAACCTCTGACACCCCTGTGGGAGCCTGGTACTCTCG	1122
Db	457	CTGAGTGTAAACACACTCCCAACCTCTGACACCTGTGGGAGCCTGGTAACTCTCG	398
QY	1123	CTACAGCTAACTTCGAGGACGACGACCTTTGATGGCGAGTATTGAGGCTTCCTTC	1182
Db	397	CTACAGCTAACTTCGAGGACGACGACCTTTGATGGCGAGTATTGAGGCTTCCTTC	338
QY	1183	CTGCTGGAGTGGAGCAGCAGTCTCGGGCTGTGAGCAGTCCAGCTGAATTCCTGCTG	1242
Db	337	CTGCTGGAGTGGAGCAGCAGTCTCGGGCTGTGAGCAGTCCAGCTGAATTCCTGCTG	278
QY	1243	GCTGCTGTGACATCTTAGCCCTGTGTTTTTGGCTCTCTTTTCTGTCAACAGCGTCCG	1302
Db	277	GCTGCTGTGACATCTTAGCCCTGTGTTTTTGGCTCTCTTTTCTGTCAACAGCGTCCG	218
QY	1303	TTCCTTGTGACATGAGAAGGACAGACAGAAAGGGAGACAAAGGGGCTGTGACTACCG	1362
Db	217	TTCCTTGTGACATGAGAAGGACAGACAGAAAGGGAGACAAAGGGGCTGTGACTACCG	158
QY	1363	CCAGCAGAGTACCGGAGCTGAGC--CTAGAGCTGATCTTGGAGAAATGTAGAG	1413
Db	157	CCAGCAGAGTACCGGAGCTGAGC--CTAGAGCTGATCTTGGAGAAATGTAGAG	98
QY	1420	--CCAGCAGAGGCACTGAGGGGGAGCGGTTACTGTCTCTCTCATTTATG--CCA	1476
Db	97	ACACAGCAGAGGCACTGAGGGGGAGCGGTTACTGTCTCTCTCATTTATG--CCA	38
QY	1477	CTTCCCTTTTAACTGCGCAAAATTTTTTAAATAAA 1512	
Db	37	CTTCCCTTTTAACTGCGCAAAANNNVTTTTTAAATAAA 2	

[illegible]

QY		511	CCCCAGCTCCGCGGGCTTCTCCTCCGAGCCCTGCACCTTGAAACTCTGGAGTTCAAGCT	572
Db		985	CSCGCCCGASTGCCCTTTKCCCGGCCGGS6CCCTGTGKA--CTCTGAKTTTCOMBT	928
QY		573	CCC GCCGCTCC CAGA CT CG GCCT GCG CA AAT TG CCACA GTGTGA ACT GA CCT GGC	632
Db		927	CCCGKGCTCCCAAGAACGCG-CTGGGCTATAATGCVACAGTGGKCATATRAACCCTCC	869
QY		633	TCTGGGCTAGAGATGAGCTCTGGAGTCCCGGAGGAGTAGTCCGGGCTCTGACAGTCACT	692
Db		868	TCTTGAGGCTATAGATGCTTGTGTCTCCGGGCGGAGTAGTACCGGTTCKGCGGCTCTCT	809
QY		693	GCACTGGGGGGCTGCAGGTCGTCCGGGCTGGAGCACATTGGAAGGCACCGTTTCCC	752
Db		808	GAACTGGGGTGTGTGACAGTCTGTCGGGCTCGGAGCACACTGTGAAGGCCAACCCTTCCC	749
QY		753	TGCGAGATCCACGTGTGTTCACTCAGACCGGCTTTGGCCAAGTTGACGAGGCTTTGGG	812
Db		748	TGCGGAGATCCACGTGTGTTCACTCAGACCGGCTTTGGCCAAGTTGAGAGGCTTTGGG	689
QY		813	GGCGCCGGAGGCGCTGGCGGTGTGGCGCGCTTTTGGAAGAGGCGCGGAAGAAACAG	872
Db		688	GGCGCCGGAGGCGCTGGCGGTGTGGCGCGCTTTTGGAAGAGGCGCGGAAGAAACAG	629
QY		873	TGCTTATGAGCATGTGCTGTCTCGCTTGGAGAAATCGCTGAGAGAGGCTCAGACATCA	932
Db		628	TGCTATATGAGCATGTGCTGTCTCGCTTGGAGAAATGCTGAGAGAGGCTTGAAGAACTCA	569
QY		933	GGTCCAGAGACATGACATATGCACTCTGACCTCTGACTTCAGCCGCTACTTTCAMTA	992
Db		568	GGTCCAGAGACATGACATATGCACTTCGACCTCTGACTTCAGCCGCTACTTTCAMTA	509
QY		993	TGAGGGGCTCTGACTACACCGGCTGTGGCCAGAGGTGATGTGACGTGTTTAACCA	1052
Db		508	TGAGGGGCTCTGACTACACCGGCTGTGGCCAGAGGTGATGTGACGTGTTTAACCA	449
QY		1053	GACAGTATGCTGAGTGTGTAAGACAGTCCACACCTCTTGACACCTGTGGGAGCTGG	1112
Db		448	AACAGTATGCTGAGTGTGTAAGACAGTCCACACCTCTTGACACCTGTGGGAGCTGG	389
QY		1113	TGACTTCGAGCTPACAGTGAATTCGAGCGACCGACCTTTGAATGGGCGAGTATTGA	1172
Db		388	TAACTCTCGGCNACAGCTGGACCTTCGAGCGACCGACCTTTGAATGGGCGAGTATTGA	329
QY		1173	GGGCTCTTCCCGAGTGGAGTGAAGACAGAGTCCCTGGGCTGTGACCCAGTCCAGCTGAA	1232
Db		328	GGGCTCTTCCCGAGTGGAGTGAAGACAGAGTCCCTGGGCTGTGACCCAGTCCAGCTGAA	269
QY		1233	TTCCTGCTGAGCTGTGTGATCATCTTACCCCTGTGTTTTGGCTCTCTTTTGTGCTGAC	1292
Db		268	TNCCTGCTGAGCTGTGTGATCATCTTACCCCTGTGTTTTGGCTCTCTTTTGTGCTGAC	209
QY		1293	CACGGTCCGCTCTTGTGACAGATAGAAAGGCACACAGAAAGGAGAAACCAAAGGGGGTGT	1352
Db		208	CAGGTCCGCTCTTGTGACAGTGTAGAAAGGCACACAGAAAGGAGAAACCAAAGGGGGTGT	149
QY		1353	GAGCTACCGCCAGCAGAGTGAAGCAGACTGAAGCTTAAGGCTGAGTCTTGAAGAAATG	1412
Db		148	GAGCTACCGCCAGCAGAGTGAAGCAGAACTGAAGGCTGAGTCTTGAAGAAATG	89
QY		1413	TGAGAAACCAAGCCAGAGCATCTTGAAGGGAGACCGGTATACGTCTCTGTCC	1462
Db		88	TGAGAAACCAAGCCAGAGCATCTTGAAGGGAGACCGGTATATGACCNAGCC	39

RESULT 8					
AL577748/C	AL577748	1031 bp	MRNA	linear	EST 01-JUN-2003
LOCUS					
DEFINITION	CDNA clone CS0DK007K10.3-PRIME, mRNA sequence.				
ACCESSION	AL577748				
VERSION	AL577748.2	GI:31316001			

	KEYWORDS	EST.
	SOURCE	Homo sapiens (human)
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	REFERENCE	Li,W.B., Gruber,C., Jeejee,J. and Polyes,D.
	AUTHORS	Full-length cDNA libraries and normalization
	TITLE	Unpublished (2001)
	JOURNAL	On Feb 16, 2001 this sequence version replaced gi:12941176.
	COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqef@genoscope.cns.fr Web : www.genoscope.cns.fr Library : seqef@genoscope.cns.fr This sequence belongs to sequence cluster 5300.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK007BF05NP&cluster=5300.f . Contact : Feng Liang Email : liang@lifeatech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DK007BF05NP.
	FEATURES	
	source	location/Qualifiers
		1..103;
		/organism="Homo sapiens"
		/mol_type="rRNA"
		/db_xref="taxon:9606"
		/clone="CS0DK007YK10"
		/cell_type="HELA CELLS COT 25-NORMALIZED"
		/clone_id="HELA"
		/note="First strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN		
Query Match	52.4%; Score 796.8; DB 9; Length 1031;	
Best Local Similarity	86.0%; Pred.No.7.9e-167;	
Matches	85; Conservative 50; Mismatches 84; Indels 5; Gaps 4;	
Oy	521 TCGCCGCTTCGCCGGCGCCTTGCGGCCTTCGAACTCGGGCTTCAGACTCCCGCGCGC	580
Dd	989 TCSCCCCNGCTCCCSCTTYRGCCCGCCCTCCSCCAATCTCTGCTSCCKTCCCGC	930
Oy	581 TCCGAATACTGGCTCGGCAACATGGACAAGTGTGCAACTGACCCTGCTCTGGCC	640
Dd	929 CYTCGMGYAATCTCTCKMAAKAGGCACA-KGTGMAAYGMCCYTSCTGCKYKGCG	871
Oy	641 TAGAGTAGCTCTGGGTCCCGGGGGAGAACCGGGGCTCGACGCTCATTC-AGCATGG	699
Dd	870 KAGAATGAGCTCTGGGTCCCGGGGGAGTAACGGGGCTTCGACGATCTMTSTKGAACKG	811
Oy	700 GGGGCTGACAGTGTGTCGGGCTCGAGACACTGTGGAAGGCCACCTTTCCCTGCGAG	759
Dd	810 GGNGCTGACAGCKGTCCAGSGTTCGAGACATGGGGGAAGGCCACCGGTGCCCGCAG	751
Oy	760 ATCAAGTGTGTTCACTTACAGACCGCTTCGACAGATTACAGAGCCTTGGGGGCGCCG	819
Dd	750 ATCAAGGGGTTCACCTACAGACCGCTTCGACAGATTACAGAGCC-TKSGGGGCGCG	692
Oy	820 GAGAGCTTGGCGGTGTGGCGGCTTCCTTGTGAGGAGGGCCCCGAAAGAAAACAGTGCNT	879
Dd	691 GAGAGCTTGGCGGTGTGGCGGCTTCCTTGTGAGGAGGGCCCCGAAAGAAAACAGGCGTAT	632
Oy	880 GAGAGTGTGTCTGCTCGCTGGAAGAAATCGCTAGGAAGGCTCAGAGACTAGGTCCA	939
Dd	631 GAGAGTGTGTCTGCTGCTGGAAGAAATCGCTAGGAAGGCTCAGAGACTAGGTCCA	572
Oy	940 GAAGTGAATATCTGCACTCTGCGCTTGTGACTTACGCCGTACTTCCAAATATAGGGG	999
Dd	571 GNACTGAAATATCTGCACTCTGCGCTTGTGACTTACGCCGTACTTCCAAATATAGGGG	512

```

QY 1000 TCTGTACATACACCGCCCTGTGACGAGGAGTGTCTCATCTGATCTGTGTTAAACAGACAGTGT 1059
DB 511 TCTGTACATACACCGCCCTGTGACGAGGAGTGTCTCATCTGATCTGTGTTAAACAGACAGTGT 452
QY 1060 ATGCTGAGTGTACAGAGCTCCACACCCCTCTCTGACACACCCCTGTGGGAGCTGTGATCTCT 1119
DB 451 ATGCTGAGTGTACAGAGCTCCACACCCCTCTCTGACACACCCCTGTGGGAGCTGTGATCTCT 392
QY 1120 CGGCTACAGCTGAACCTCCAGACGACGACCTTTGATGGGAGGATGATGATGAGGCTCC 1179
DB 391 CGGCTACAGCTGAACCTCCAGACGACGACCTTTGATGAGGAGGATGATGATGAGGCTCC 332
QY 1180 TTCCCTGTCTGAGTGTGACAGCAGCTCTGAGGCTGTGAGCCAGTCCAGCTGATCTCTG 1239
DB 331 TTCCCTGTCTGAGTGTGACAGCAGCTCTGAGGCTGTGAGCCAGTCCAGCTGATCTCTG 272
QY 1240 CTGGCTGTCTGAGTGTGACATCTCTAGCCTGAGTGTGAGCTCTCTTTTGTGCTGACAGGCTC 1299
DB 271 CTGGCTGTCTGAGTGTGACATCTCTAGCCTGAGTGTGAGCTCTCTTTTGTGCTGACAGGCTC 212
QY 1300 GCGTTCCTTGTGAGTGTGAGAAAGAGACAGACAGAGGAGAACCAAGGGGCTGTGAGCTAC 1359
DB 211 GCGTTCCTTGTGAGTGTGAGAAAGAGAGACAGACAGAGGAGAACCAAGGGGCTGTGAGCTAC 152
QY 1360 CGGCCAGAGAGGTGTACCGAGACTGAGAGCTGAGAGCTGAGATCTTGTGAGAAATGTGAGAG 1419
DB 151 CGGCCAGAGAGGTGTACCGAGACTGAGAGCTGAGAGCTGAGATCTTGTGAGAAATGTGAGAG 92
QY 1420 CCAGCCAGAGAGCTGTGAGAGGAGAGCCGATACCTGTCTGTCTGATATGAGCA--C 1477
DB 91 CCAGCCAGAGAGCTGTGAGAGGAGAGCCGATATCTGTCTGTCTGATATGAGCA--C 32
QY 1478 TTCTTTTAACTGCGCAAGAAATTTTAAAA 1508
DB 31 TCTATTTTAACTGCGCAAGAAATTTTAAAA 1

```

```

RESULT 9
AL580216/c 1017 bp mRNA linear EST 01-JUN-2003
LOCUS AL580216 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0D005YK14 3-PRIME, mRNA sequence.
ACCESSION AL580216
VERSION AL580216.2 GI:31318495
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1017)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Poljansky, D.
TITLE Full-length cDNA libraries and normalization.
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12946023.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5300.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D005BFO7NP1&cluster=5300.f. Contact:
Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CS0D005BFO7NP1.
Location/Qualifiers
1..1017
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D005YK14"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

```

```

/origin="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

Query Match 51.9%; Score 789.4; DB: 9; Length 1017;
Best local similarity 87.9%; Pred. No. 3.5e-165;
Matches 832; Conservative 44; Mismatches 66; Indels 4; Gaps 3;

```

```

QY 545 GCGCCCTGGAACCTCTGAGGCTCCAGCTCCGCGCTCCAGAACTCCGCTCCGCAACA 604
DB 943 GCGCCCTGGAACCTCTGAGG--TTCCAGTCCGCGCTCCAGAACTCCGCTCCGCAACA 886
QY 605 ATGGCCAGTGTGCAACTGACCTGCTT-CTTGGCTGAGATGAGCTCTGGGTCCGGG 663
DB 885 ATGGCCAGTGTGCAACTGACCTGCTTCTCCGCTGAGATGAGCTCTGGGTCCGGG 826
QY 664 CGGAGTACCGGGCTCTGAGCTGATCTGACCTGAGGAGCTGACAGTCTCCGAGCTCG 723
DB 825 CGGAGTACCGGGCTCTGAGCTGATCTGACCTGAGGAGCTGACAGTCTCCGAGCTCG 766
QY 724 GAGCACTGTGAGAGGCAACCTTTCTGCGGAGATCCAGTGTTCATCTCAGCACC 783
DB 765 GAGCACTGTGAGAGGCAACCTTTCTGCGGAGATCCAGTGTTCATCTCAGCACC 706
QY 784 GCTTTGCAAGTGTGACAGGCTTGGGAGGCGCGGAGAGCTGAGCGGTGGGCGCG 843
DB 705 VCGAAACCCAGATGTGAGAGGCTTGGGAGGCGCGGAGAGCTGAGCGGTGGGCGCG 646
QY 844 TTTCTGAGAGGAGGCGCGGAGAGAGAGTGTCTGAGAGAGTGTCTGTCTGAGAA 903
DB 645 TTTCTGAGAGGAGGCGCGGAGAGAGAGTGTCTGAGAGAGTGTCTGTCTGAGAA 586
QY 904 GAAATCGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 963
DB 585 GAAATCGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 526
QY 964 CCTCTGACTCAGCGCTTACTTCCAAATATGAGGAGTCTGTGACTACAGCCCTGTG 1023
DB 525 CCTCTGACTCAGCGCTTACTTCCAAATATGAGGAGTCTGTGACTACAGCCCTGTG 466
QY 1024 CAGGCTGTGATGTGAGTGTGATTTAAACAGAGTGTGAGTGTGAGTGTGAGTGTGAG 1083
DB 465 CAGGCTGTGATGTGAGTGTGATTTAAACAGAGTGTGAGTGTGAGTGTGAGTGTGAG 406
QY 1084 ACCCTCTGAGACACCTGTGAGGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1143
DB 405 ACCCTCTGAGACACCTGTGAGGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 346
QY 1144 ACGCAGCTTTGATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 1203
DB 345 ACGCAGCTTTGATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 286
QY 1204 CCTCGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 1263
DB 285 CCTCGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 226
QY 1264 CTGGTGTGAGGCTCTCTTTTGTGCTGACAGGCTGTGAGGCTGTGAGGAGTGTGAGGAG 1323
DB 225 CTGGTGTGAGGCTCTCTTTTGTGCTGACAGGCTGTGAGGCTGTGAGGAGTGTGAGGAG 166
QY 1324 CGCAGCAGAGGAGGAGACCAAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1383
DB 165 CGCAGCAGAGGAGGAGACCAAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 107
QY 1384 GAGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1443
DB 106 GAGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 47

```

QY	1444	GGCGGAACTGCTCGTCTGTGATATTCGACATCTTTTACT	1489
D5	46	GGCGGTGNGNGKGTGGCCCTGSMNATNATGCAAGTCTTTTACT	1
RESULT 10			
LOCUS	AL555184		
DEFINITION	AL555184 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens	1067 bp	trNA linear EST 31-MAY-2003
ACCESSION	AL555184		
VERSION	AL555184.2	GI:31276993	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	L.I.W.B., Gruber,C., Jesse,J. and Polayes,D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12896674.		

FEATURES

Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5300.f For
more information about this cluster, see
[http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK007BF05Q51&cluster=5300.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK007BF05Q51&cluster=5300.f). Contact :
Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/invitrogen/Corporation/1600>
Faraday Avenue Genoscope sequence ID : CS0DK007BF05Q51.
Location/Qualifiers

```

source
1..1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDK007YK10"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR I
sites of the pCMVSPORT 6 vector. Library was normalized

```

Query Match	50.8%	Score 772.6	DB 9	Length 1067
Best Local Similarity	96.0%	Pred. 26-161		
Matches 840	Conservative 5	Mismatches 24	Indels 6	Gaps 5
QY	1	ACAGTCAGCCGCAATGCTCCCTCTGTGCCCCAGACCCCTTGTGCTCTCTGTGTGATCCCGGCC	60	
Db	78	ACAAGTCAGCCGCAATG-TCCCTGTGCCACCCCTGTGCTCTCTGTGTGATCCCGGCC	136	
QY	61	CCGCTCAGAGCCTCACTGTGCACATGCTGCTCATGCTGCTTGTGATGCTGTGCAT	120	
Db	137	CTGTCTCAGAGCCTCACTGTGCACATGCTGCTGTACATGCTGTTCGTGTGCTGTGCAT	196	
QY	121	CCCCAGAGTTGCCCCCGGATGCGAGAGATTTCCCTTTGGAGAGAGCTCTTCTGGGGAA	180	
Db	197	CCCCAGAGTTGCCCCCGGATGCGAGAGATTTCCCTTTGGAGAGAGCTCTTCTGGGGAA	256	
QY	181	GATGACCCACTGGGCGAGAGATCTGCCCATGTGAAGAGATTCACCACAGAGAGAGAT	240	
Db	257	GATGACCCACTGGGCGAGAGATCTGCCCATGTGAAGAGATTCACCACAGAGAGAGAT	316	
QY	241	CCACCGGAGAGAGATCTCACTGTGAAGAGAGATCTCACTGTGAAGAGAGATCTCACT	300	
Db	317	CCACCGGAGAGAGATCTCACTGTGAAGAGAGATCTCACTGTGAAGAGAGATCTCACT	376	

QY	301	GAAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGMACTTGAGAGATCTACTACTTT	360
Db	377	GAAAGTTAAGCTTAATCAGAAAGAGAGGCTCCCTGMACTTGAGAGATCTACTACTTT	436
QY	361	GAGGCTCTCGAGAGATCTCTAAAGAACCCAGAAATAATAGCCACAGGAGCAAGAAAGGGAT	420
Db	437	GAGGCTCTCGAGAGATCTCTCAAGAACCCAGAAATAATAGCCACAGGAGCAAGAAAGGGAT	496
QY	421	GACCAGAGTCATTGGCGCTATGAGAGCGACCGGCTTGCGCCCGGGTGTGCCAGCTGC	480
Db	497	GACCAGAGTCATTGGCGCTATGAGAGCGACCGGCTTGCGCCCGGGTGTGCCAGCTGC	556
QY	481	GGGGGCGGCTTTCAGATCCCGCGGTGGAATCCGGGCCAGCTGGCGGCTTTCGGCCGGGCG	540
Db	557	GGGGGCGGCTTTCAGATCCCGCGGTGGAATCCGGGCCAGCTGGCGGCTTTCGGCCGGGCG	616
QY	541	CTGCGGCCCTTGGAAGCTCTGGGCTTTCAGACTCCGCGCTCCCAAACTGCGCTGCGC	600
Db	617	CTGCGGCCCTTGGAAGCTCTGGGCTTTCAGACTCCGCGCTCCCAAACTGCGCTGCGC	676
QY	601	AACATGCGACAGTGTGAGCACTGACCCGTCGCTCGGCTGTAGATGTGCTGAGCTCC	660
Db	677	AACATGCGACAGTGTGAGCACTGACCCGCTCTCTTGAGGTAGAGATGTGCTGAGCTCC	736
QY	661	GGGCGGGAGTACCGGGCTCTGACAGCTCATCTGCACT - GGGGGGCTGCAGATCGTCCGAG	719
Db	737	GGGCGGGAGTACCGGGCTCTGACAGCTCATCTGCACTGAGGGGGGCTGCAGATCGTCCGAG	796
QY	720	CTTGGAGGCACTGTGTGGAAGGCCACCGTTCCTCGCGAGATCTCAAGTGTGTCACCTGAG	779
Db	797	CTTGGAGGCACTGTGTGGAAGGCCACCGTTCCTCGCGAGATCTCAAGTGTGTCACCTGAG	855
QY	780	CACCGGCTTTCAGAGTTGACGAGGCGTTGGGGCGGCCGGAGAGGCTGCGCTGTGTGCG	839
Db	856	CACAGCTTTCAGAGTTGACGAGGCGTTGGGGCGGCCGGAGAGGCTGCGCTGTGTGCG	912
QY	840	CGGCTTTCGAGAGAGGGCGCCGGAAGAAAACAAGG	874
Db	913	GCCTTTCGAGAGGGCGCCGGAAGAAAACAAGTCTATG	947

```

RESULT 11
BG386425
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BG386425      874 bp      mRNA      linear      EST 12-MAR-2001
602455652F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4583869 5',
mRNA sequence.
BG386425
BG386425.1   GI:13279871
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 874)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LICM1308 row: h column: 14
High quality sequence stop: 714.
Location/Qualifiers
1..874
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```


QY 909 CGCTGAGGAGGCTCAGAGACTCAGGTCGCCAGAGCTGACATATCTGACTCTCCCTC 968
 Db 629 CGCTGAGGAGGCTCAGAGACTCAGGTCGCCAGAGCTGACATATCTGACTCTCCCTC 570
 QY 969 TGACTTACGCGCTACTCTCCAAATATGAGGGGCTCTGATTAACCCGCTCTGCGCCAGG 1028
 Db 569 TGACTTACGCGCTACTCTCCAAATATGAGGGGCTCTGATTAACCCGCTCTGCGCCAGG 510
 QY 1029 TGCTATCTGACTGCTGTTTAAACAGACAGTATGCTGATGCTTAAGACAGCTCCACACCT 1088
 Db 509 TGCTATCTGACTGCTGTTTAAACAGACAGTATGCTGATGCTTAAGACAGCTCCACACCT 450
 QY 1089 CTCTGACACCCCTGTGGGACCTGCTGATCTCGGCTACAGCTGAACTTCCAGACGCA 1148
 Db 449 CTCTGACACCCCTGTGGGACCTGCTGATCTCGGCTACAGCTGAACTTCCAGACGCA 390
 QY 1149 GCCTTGAATGGGAGAGTATGAGGCTCTCTCCCTGCTGAGTGAAGACAGAGCTCTCG 1208
 Db 389 GCCTTGAATGGGAGAGTATGAGGCTCTCTCCCTGCTGAGTGAAGACAGAGCTCTCG 330
 QY 1209 GCGTGTGAGCCAGCTCAGCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1268
 Db 329 GCGTGTGAGCCAGCTCAGCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 270
 QY 1269 TTTTGGCTCTCTTTTGTCTGCTCAACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1328
 Db 269 TTTTGGCTCTCTTTTGTCTGCTCAACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 210
 QY 1329 CAGAGGGGAGACCAAGGGGGGTGTGAGCTACCGCCAGCAGAGTGTAGCCGAGCTGAGC 1388
 Db 209 CAGAGGGGAGACCAAGGGGGGTGTGAGCTACCGCCAGCAGAGTGTAGCCGAGCTGAGC 150
 QY 1389 CTGAGGCTGAGATCTTGAAGAAATGTGAAGAGCAGCAGAGCACTTAAGGGGAGCCG 1448
 Db 149 CTGAGGCTGAGATCTTGAAGAAATGTGAAGAGCAGCAGAGCACTTAAGGGGAGCCG 90
 QY 1449 TAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1508
 Db 89 TAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 30
 QY 1509 TAAATATTTTAAAT 1522
 Db 29 TAAATATTTTAAAT 16

RESULT 13
 A1831707/c 682 bp mRNA linear EST 21-DEC-1999
 LOCUS A1831707
 DEFINITION wj40h01.x1 NCI CGAP Lul9 Homo sapiens cDNA clone IMAGE:2405329 3'
 similar to TR:Q16790 Q16790 CARBONIC ANHYDRASE 1, mRNA sequence.
 A1831707
 ACCESSION A1831707.1 GI:5452464
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 682)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: gsp@nci.nih.gov
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 www-bio.1nl.gov/dbirp/image/image.html

Insert Length: 1322 Std Error: 0.00
 Seq primer: -40UP from Gldco
 High quality sequence stop: 424.
 Location/Qualifiers
 1..682
 source

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2405329"
 /tissue_type="squamous cell carcinoma, poorly
 differentiated (4 pooled tumors, including primary and
 metastatic)"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI-CGAP_Lul9"
 /note="Organ: Lung; Vector: pRT3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 pooled lung tumor tissue, and was then primed with a Not I
 - oligo(dT) primer. Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pRT3D vector. Library went through one round of
 normalization. Library constructed by Bento Soares and M.
 Fatima Bonaldo."

ORIGIN

Query Match 43.7%; Score 665.8; DB 9; Length 682;
 Best Local Similarity 98.8%; Pred. No. 1.1e-137;
 Matches 670; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 845 TTCTGAGAGAGGCGCCGAGAGAAACAGTGCCTATGACAGCTGCTGCTGCTGAG 904
 Db 678 TTTTGAAGAGAGGCGGAGAGAAACAGTGCCTATGACAGCTGCTGCTGCTGAG 619
 QY 905 AATGCTGAGAGAGGCTCAGAGCTCAGGCTCCAGAGCTGAGATATCGCACTCTG 964
 Db 618 ACATGCTGAGAGAGGCTCAGAGCTCAGGCTCCAGAGCTGAGATATCGCACTCTG 559
 QY 965 CCTGACTTCAAGCGCTACTTCCAAATATGAGGGGCTCTGACTACACCGCCCTGTCCC 1024
 Db 558 CTTTGACTTCAAGCGCTACTTCCAAATATGAGGGGCTCTGACTACACCGCCCTGTCCC 499
 QY 1025 AGGCTGTACTTGTGACTGTGTTTAAACAGAGAGAGTGTAGTGTAGAGAGCTCCACA 1084
 Db 498 AGGCTGTACTTGTGACTGTGTTTAAACAGAGAGAGTGTAGTGTAGAGAGCTCCACA 439
 QY 1085 CCTCTGTGACACCTGTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1144
 Db 438 CCTCTGTGACACCTGTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
 QY 1145 CGAGCTTGAATGGGAGAGTGAAGGCTCTCTCTGCTGAGTGAACAGCACTC 1204
 Db 378 CGAGCTTGAATGGGAGAGTGAAGGCTCTCTCTGCTGAGTGAACAGCACTC 319
 QY 1205 CTGAGGCTGCTGAGCAGTCCAGTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1264
 Db 318 CTGAGGCTGCTGAGCAGTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
 QY 1265 TGGTTTGGCTCTCTTTTGTCTGCTCAGCAGGCTGCTCTCTGCTGAGTGAAGAGC 1324
 Db 258 TGGTTTGGCTCTCTTTTGTCTGCTCAGCAGGCTGCTCTCTGCTGAGTGAAGAGC 199
 QY 1325 AGCAGAGAGAGGAGCAAAAGGGGGTGTGAGTACCGCCAGCAGAGAGTGAAGCAGAGCTG 1384
 Db 198 AGCAGAGAGAGGAGCAAAAGGGGGTGTGAGTACCGCCAGCAGAGAGTGAAGCAGAGCTG 139
 QY 1385 GAGCTTGAAGGCTGATCTTGAAGATGTGAAGAGCAGCAGAGCAGTCTGAGGGGAG 1444
 Db 138 GAGCTTGAAGGCTGATCTTGAAGATGTGAAGAGCAGCAGAGCAGTCTGAGGGGAG 79
 QY 1445 CCGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1504
 Db 78 CCGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19

QY 1505 AAAATAAATTATTAAT 1522
 DB 18 AAAATAAATTATTAAT 1

RESULT 14
 BX383092 1074 bp mRNA linear EST 08-MAY-2003
 LOCUS BX383092 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0DJ005YG10 5-PRIME, mRNA sequence.
 ACCESSION BX383092 GI:30449113
 VERSION EX383092.1 GI:30449113
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5300.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ005BD05GPI&cluster=5300.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DJ005BD05GPI.
 Location/Qualifiers
 1..1074
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ005YG10"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_idb="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-clisio(NT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 43.7%; Score 665.4; DB 13; Length 1074;
 Best Local Similarity 93.1%; Pred. No. 1.7e-137;
 Matches 734; Conservative 14; Mismatches 6; Indels 34; Gaps 4;

QY 1 ACAGTACGCGCATGCTCCCTGTCGCCACCCCTGAGCTCCCTTTATCCCGCC 60
 DB 70 ACAGTACGCGCATGCTCCCTGTCGCCACCCCTGAGCTCCCTTTATCCCGCC 129

QY 61 CCTGCTCCAGGCTCACTGTGCACTGCTGTCTCACTGCTGCTTGTATGCTGTCAT 120
 DB 130 CCTGCTCCAGGCTCACTGTGCACTGCTGTCTCACTGCTGCTTGTATGCTGTCAT 189

QY 121 CCCGAGAGGTTGCCCCCGGATGCGAGAGATTCCTCCCTTGGAGAGAGGCTCTTGGGGAA 180
 DB 190 CCCGAGAGGTTGCCCCCGGATGCGAGAGATTCCTCCCTTGGAGAGAGGCTCTTGGGGAA 249

QY 181 GATGACCCACTGGGCGAGAGAGATCTGCCAAGTGAAGAGATTACCCAGAGAGAGAT 240
 DB 250 GATGACCCACTGGGCGAGAGAGATCTGCCAAGTGAAGAGATTACCCAGAGAGAGAT 309

QY 241 CCACCCGAGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATTAAGT 300
 DB 310 CCACCCGAGAGAGATCTACCTGAGAGAGAGATCTACCTGAGAGAGAGATTAAGT 369

QY 301 GAAGTTAAGCTTAAATCAGAGAGAGGCTCCCTGAAGTTAGAGATCTACTGTT 360
 DB 370 GAAGTTAAGCTTAAATCAGAGAGAGGCTCCCTGAAGTTAGAGATCTACTGTT 429

QY 361 GAGGCTCTGAGAGCTCTCAGAGAACCCCAATTAATGCCACAGGAGACAAAGAGGAGAT 420
 DB 430 GAGGCTCTGAGAGCTCTCAGAGAACCCCAATTAATGCCACAGGAGACAAAGAGAT 482

QY 421 GACCAGAGTATTGGCGCTATGAGAGGACCCGCTGAGCCCGGGTGTCCCCAGCCTGC 480
 DB 483 -----AGCGAGACCCGCTGTGCCCCGGTGTCCCCAGCCTGC 519

QY 481 GCGGAGCCGCTTCAGTCCCGGCTGATATCCGCGCCACAGCTGCGGCTTCTGCCGCGC 540
 DB 520 GCGGAGCCGCTTCAGTCCCGGCTGATATCCGCGCCACAGCTGCGGCTTCTGCCGCGC 579

QY 541 CTGCGCCCTCTGAACTCTGAGCTTCCAGCTTCCGCGCTCCAGAGACTGCGGCTGCGC 600
 DB 580 CTGCGCCCTCTGAACTCTGAGCTTCCAGCTTCCGCGCTCCAGAGACTGCGGCTGCGC 639

QY 601 AACATGAGCAAGGTGAGCACTGACCCCTGCTGAGCTAGAGATGCTGAGTCC 660
 DB 640 AACATGAGCAAGGTGAGCACTGACCCCTGCTGAGCTAGAGATGCTGAGTCC 699

QY 661 GGGCGGAGTACCGGGCTCTGACAGCTGACATCTGACCTGAGGGGCTGACAGTGTGTCGGCC 720
 DB 700 GGGCGGAGTACCGGGCTCTGACAGCTGACATCTGACCTGAGGGGCTGACAGTGTGTCGGCC 759

QY 721 TCGAGAGCAACTGTGAGAGAGCCACCGTTTCCCGGAGATCAGCTGATCACTCAGC 780
 DB 760 TC-GGRSMACTGTGAGAGAGCACS--TTTCTTCCGAGAT-CMGCTGTTTCACTCAGC 815

QY 781 ACCGCTT 788
 DB 816 ACGCTTT 823

RESULT 15
 BU620600/c 691 bp mRNA linear EST 23-SEP-2002
 LOCUS BU620600 UI-H-FL1-bfX-o-01-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
 DEFINITION UI-H-FL1-bfX-o-01-0-UI 3', mRNA sequence.
 ACCESSION BU620600
 VERSION BU620600.1 GI:23286815
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: James Martin
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 The following repetitive elements were found in this cDNA
 sequence: 1-45, >AT rich#low_complexity (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
 source location/Qualifiers
 1..691
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FL1-bfX-o-01-0-UI"

/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FL1"
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac
(Pharmacia) with a modified polylinker; Site 1: Ecor I;
Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library
derived from a pool of mRNA obtained from 4 cell lines
from grade III chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
GAGTGGGTG. The cell lines were provided by Dr. James
Martin from the University of Iowa.
TAG_TISSUE=human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FL1
TAG_SEQ=GAGTGGGTG"

ORIGIN

Query Match 43.4%; Score 660.4; DB 13; Length 691;
Best Local Similarity 99.7%; Pred. No. 1.8e-136;
Matches 672; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 849 GGAGGAGGCGCCGGAAGAAACAGTGCCTATGAGCAGTGTCTCTGCTTGGAGAAAT 908
DB 691 GGAGGAGGCGCCGGAAGAAACAGTGCCTATGAGCAGTGTCTCTGCTTGGAGAAAT 632
QY 909 CGCTGAGGAGGCTCAGAGACTAGGTCCAGAGCTGAGCATATCTGCACTCCGCCCTC 968
DB 631 CGCTGAGGAGGCTCAGAGACTAGGTCCAGAGCTGAGCATATCTGCACTCCGCCCTC 572
QY 969 TGACTCAGCGGCTACTTCCATATAGAGGGGTCTCTGACTACACCGCCCTGTGCCAGG 1028
DB 571 TGACTCAGCGGCTACTTCCATATAGAGGGGTCTCTGACTACACCGCCCTGTGCCAGG 512
QY 1029 TGTCTATGAGTGTGTTTAAACAGACAGTATGATGCTGATGCTTAAAGCAGCTCCACCT 1088
DB 511 TGTCTATGAGTGTGTTTAAACAGACAGTATGATGCTGATGCTTAAAGCAGCTCCACCT 452
QY 1089 CTCTGACACCCCTGTGGGAGCTGTGATCTCGGGCTACAGCTGAACCTCCAGCGACGCA 1148
DB 451 CTCTGACACCCCTGTGGGAGCTGTGATCTCGGGCTACAGCTGAACCTCCAGCGACGCA 392
QY 1149 GCCTTTGAATGGGAGATGATTGAGGCTTCCTTCCCTGCTGAGTGAAGACAGTCTCTG 1208
DB 391 GCCTTTGAATGGGAGATGATTGAGGCTTCCTTCCCTGCTGAGTGAAGACAGTCTCTG 332
QY 1209 GGCTGCTGAGCAAGTCCAGTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1268
DB 331 GGCTGCTGAGCAAGTCCAGTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 273
QY 1269 TTTTGGCCTCCTTTTGTGCTGACAGCGTGCCTTCTTGTGCAATGAGAGGACACA 1328
DB 272 TTTTGGCCTCCTTTTGTGCTGACAGCGTGCCTTCTTGTGCAATGAGAGGACACA 213
QY 1329 CAGAGGAGGAGCAACAAAGGGGGTGTGAGCTACCGCCAGAGAGGTAGCCGAGACTGAGAGC 1388
DB 212 CAGAGGAGGAGCAACAAAGGGGGTGTGAGCTACCGCCAGAGAGGTAGCCGAGACTGAGAGC 153
QY 1389 CTAGAGGCTGATCTTGGAGATGTGAGAGCCAGCCAGAGGATCTGAGGGGAGCCGG 1448
DB 152 CTAGAGGCTGATCTTGGAGATGTGAGAGCCAGCCAGAGGATCTGAGGGGAGCCGG 93
QY 1449 TAACTGCTGCTCTGCTCATTTATGCACTTCTTTTAACTGCAAGAAATTTTAA 1508
DB 92 TAACTGCTGCTCTGCTCATTTATGCACTTCTTTTAACTGCAAGAAATTTTAA 33

QY 1509 TAAATATTTATAT 1522
DB 32 TAAATATTTATAT 19

Search completed: March 7, 2004, 05:23:47
Job time: 2834 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 5, 2004, 08:58:24 / Search time 59 Seconds
(without alignments)
2198.124 Million cell updates/sec

Title: US-09-967-237-2

Perfect score: 2424
Sequence: 1 MAPLCPSPMLPILIPAPAPQ.....RRGTGVSYPRAVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq 29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1980s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2424	100.0	459	2	AAR8058 Protein e
2	2424	100.0	459	3	AAV53228 Human MN
3	2424	100.0	459	3	AAAB03005 Human MN
4	2424	100.0	459	5	AAE17175 Human RCC
5	2424	100.0	459	6	ABR58596 Human can
6	2424	100.0	459	6	ABP97744 Antino aci
7	2424	100.0	459	6	ABU56556 Lung can
8	2424	100.0	459	6	ABU56408 Lung can
9	2419	99.8	459	6	ABR52848 Kidney ca
10	2007	82.8	377	3	AAV53245 MN protei
11	2007	82.8	377	3	AAAB03021 Mature hu
12	1398.5	57.7	429	2	AAAR41746 MN protei
13	1370	56.5	257	3	AAV53241 MN protei
14	1370	56.5	257	3	AAAB03018 Human MN
15	1370	56.5	257	3	AAAR97235 Muru puta
16	1374	56.3	256	2	AAAR97235 Human PRO
17	562	23.2	337	4	AAAB66209 Human PRO
18	562	23.2	337	4	AAU29157 Human PRO
19	562	23.2	337	4	AAAB87562 Human PRO
20	562	23.2	337	5	ABG95887 Human sec
21	562	23.2	337	6	ABU58533 Human PRO
22	562	23.2	337	6	ABU88081 Novel hum
23	562	23.2	337	6	ABU84396 Human sec
24	562	23.2	337	6	ABR66270 Human sec
25	562	23.2	337	6	ABR65660 Human sec

26	562	23.2	337	6	ABU99600 Human sec
27	562	23.2	337	6	ABU82839 Human PRO
28	562	23.2	337	6	ABU89960 Novel hum
29	562	23.2	337	6	ABR68209 Human sec
30	562	23.2	337	6	ABU96262 Novel hum
31	562	23.2	337	6	ABU92693 Human sec
32	562	23.2	337	6	ABO08770 Human sec
33	562	23.2	337	6	ABO02822 Human sec
34	562	23.2	337	6	ABR74976 Human sec
35	562	23.2	337	6	ABR94738 Human PRO
36	562	23.2	337	6	ABU85711 Human PRO
37	562	23.2	337	6	ABU98871 Novel hum
38	562	23.2	337	6	ABU98086 Novel hum
39	562	23.2	337	6	ABU91792 Novel hum
40	562	23.2	337	6	ABU89485 Human PRO
41	562	23.2	337	6	ABU86326 Human sec
42	562	23.2	337	6	ABU67539 Human sec
43	562	23.2	337	6	ABU80567 Human PRO
44	562	23.2	337	6	ABU90912 Novel hum
45	562	23.2	337	6	ABO33970 Human sec

ALIGNMENTS

RESULT 1
AAR8058
ID AAR8058 standard; protein; 459 AA.

AC AAR8058;
DT 25-MAR-2003 (revised)
DT 25-JUL-1996 (first entry)

DE Protein encoded by Mutu putative oncogene MN.

KM Mutu; endogenous; cellular component; MN; Hela cell; diagnosis;
KM Lymphocytic choriomeningitis virus; LCMV; putative oncogene; treatment;
KM neoplastic; pre-neoplastic; disease; antisense therapy; antibody;
KM vaccine; vertebrate; immunisation; carbonic anhydrase.

OS Homo sapiens.

XX	Key	Location/Qualifiers
FT	Peptide	1..37
FT	Peptide	/label= sig_peptide
FT	Peptide	36..51
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Region	38..135
FT	Peptide	/note= "region homologous to collagen alpha 1 chain"
FT	Peptide	55..60
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Peptide	62..67
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Peptide	68..91
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Peptide	127..147
FT	Domain	/note= "anti-MN antibody epitope"
FT	Peptide	136..139
FT	Peptide	/note= "carbonic anhydrase domain"
FT	Peptide	279..291
FT	Peptide	/note= "anti-MN antibody epitope"
FT	Region	414..433
FT	Region	/note= "intracellular transmembrane region"
FT	Region	434..459
FT	Region	/note= "intracellular C-terminus"
FT	Peptide	435..450
FT	Peptide	/note= "anti-MN antibody epitope"
XX	MO9534650-A2.	
XX	21-DEC-1995.	

PF 15-JUN-1995; 95WO-US007628.
 XX
 PR 15-JUN-1994; 94US-00260190.
 PR 07-JUN-1995; 95US-00477504.
 PR 07-JUN-1995; 95US-00481658.
 PR 07-JUN-1995; 95US-00485049.
 PR 07-JUN-1995; 95US-00485862.
 PR 07-JUN-1995; 95US-00485863.
 PR 07-JUN-1995; 95US-00486756.
 PR 07-JUN-1995; 95US-00487077.
 XX
 PA (CIBA) CIBA CORNING DIAGNOSTICS CORP.
 PA (VIRO-) INST VIROLOGY.
 XX
 FI Zavađa J, Pastorekova S, Pastorek J;
 XX
 DR WPI; 1996-049679/05.
 DR N-PSDB; AAT09186.
 XX
 PT MN gene, protein and nucleic acid fragments - used as primers and probes
 PT in the detection of MN antigens and antibodies, and in the treatment of
 PT (pre)neoplastic disease.
 XX
 PS Claim 12; Fig 1; 102pp; English.
 XX
 CC The present sequence is encoded by the full length Mutu endogenous
 CC cellular component, MN, cDNA clone, which was isolated from lymphocytic
 CC choriomeningitis virus (ICMV) infected HeLa cells. Persistent ICMV, the
 CC exogenous Mutu transmissible agent (MX), infection increases the
 CC expression level of the MN gene. MN is a putative oncogene, and can
 CC therefore be used in the development of prods. for the diagnosis and
 CC treatment of neoplastic (NP), or pre-NP diseases. NP diseases can be
 CC treated using DNA antisense to MN transcribed mRNA, anti-MN protein
 CC antibodies can be used for the diagnosis NP or pre-NP diseases and a
 CC vaccine cong. immunogenic amounts of the MN protein can be used to
 CC immunise a vertebrate against a NP disease associated with MN antigen
 CC expression. (Updated on 25-MAR-2003 to correct PR field.)
 CC
 XX
 SQ Sequence 459 AA;

Query Match 100.0%; Score 2424; DB 2; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1,1e-182;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPFWLPLIPAPAGLTVQLLISLILMPVHPQLPKQEBSPICGGSSGSDPL 60
 DB 1 MAPLCSPFWLPLIPAPAGLTVQLLISLILMPVHPQLPKQEBSPICGGSSGSDPL 60
 QY 61 GEEDLPSEEDSPREDDPGEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEEDSPREDDPGEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPOEPONNAHDKEDQSHMRVGDDPMPWRVSPACAGRFQSVDIRPOLAFCPLRPL 180
 DB 121 DPOEPONNAHDKEDQSHMRVGDDPMPWRVSPACAGRFQSVDIRPOLAFCPLRPL 180
 QY 181 ELIGFOLPPLPELRLNNGSVOLTLPPEGLMALGPKEERFALQQLHLHGAAGRPSEHT 240
 DB 181 ELIGFOLPPLPELRLNNGSVOLTLPPEGLMALGPKEERFALQQLHLHGAAGRPSEHT 240
 QY 241 VEGHRFPAEIHVHLSTAFARVDEALGRPGGLAVLAFLKEEGEENSAYEQILSRLEBTA 300
 DB 241 VEGHRFPAEIHVHLSTAFARVDEALGRPGGLAVLAFLKEEGEENSAYEQILSRLEBTA 300
 QY 301 BEGSETQVPGDISALPSPDSRFQYEGSLTPPCAGQYIVTFVQTLWSAKQHLTIS 360
 DB 301 BEGSETQVPGDISALPSPDSRFQYEGSLTPPCAGQYIVTFVQTLWSAKQHLTIS 360
 QY 361 DTWGPDSRLQLNFRATQPLNGRVIAESPAGVSSPRAAEVQVQNSCLAAGDIALVVF 420
 DB 361 DTWGPDSRLQLNFRATQPLNGRVIAESPAGVSSPRAAEVQVQNSCLAAGDIALVVF 420
 QY 421 GULFAVTSVAFVQMRQRHRTGKGVSYRPAEVAERGA 459

DB 421 GULFAVTSVAFVQMRQRHRTGKGVSYRPAEVAERGA 459

RESULT 2
 AAY53228
 ID AAY53228 standard; protein; 459 AA.
 XX
 AC AAY53228;
 XX
 DT 16-JUN-2000 (first entry)
 XX
 DE Human MN protein SEQ ID NO:2.
 XX
 KM Human; MN protein; MN gene; oncogene; carbonic anhydrase; tumour;
 KM oncogenesis; diagnosis; neoplastic disease; cancer; carcinoma;
 KM MN/CA IX isoenzyme.
 XX
 OS Homo sapiens.
 XX
 PN US6027887-A.
 XX
 PD 22-FEB-2000.
 XX
 PF 24-JAN-1997; 97US-00787739.
 XX
 PR 21-OCT-1992; 92US-00964589.
 PR 30-DEC-1993; 93US-00177093.
 PR 15-JUN-1994; 94US-00260190.
 PR 07-JUN-1995; 95US-00477504.
 PR 07-JUN-1995; 95US-00481658.
 PR 07-JUN-1995; 95US-00485049.
 PR 07-JUN-1995; 95US-00485862.
 PR 07-JUN-1995; 95US-00485863.
 PR 07-JUN-1995; 95US-00486756.
 PR 07-JUN-1995; 95US-00487077.
 XX
 PA (SLSC-) SLOVAK ACAD SCI INST VIROLOGY.
 XX
 FI Pastorek J, Zavađa J, Pastorekova S;
 XX
 DR WPI; 2000-194827/17.
 DR N-PSDB; AAA16540.
 XX
 PT Nucleic acid based assay for diagnosing a wide variety of
 PT preneoplastic/neoplastic disease comprises screening for the presence of
 PT abnormal MN gene expression in a vertebrate.
 XX
 PS Disclosure; Fig 1; 87pp; English.
 XX
 CC The present invention describes a method of screening for
 CC preneoplastic/neoplastic disease. The method comprises: (1) determining
 CC whether abnormal MN gene expression is present in a vertebrate; and (2)
 CC if abnormal MN gene expression is determined to be present in the
 CC vertebrate, determining that the vertebrate has a significant risk of
 CC having preneoplastic/neoplastic disease. The MN gene is an oncogene and
 CC encodes an MN protein (also referred to as MN/CA IX isoenzyme). The MN
 CC protein is a tumour associated carbonic anhydrase isoenzyme. The method
 CC is used for detecting a wide variety of preneoplastic/neoplastic diseases
 CC in a vertebrate, preferably a human. The disease detected is mammary,
 CC bladder, renal, urinary tract, ovarian, uterine, cervical, endometrial,
 CC vaginal, vulval, prostate, liver, lung, skin, thyroid, pancreatic,
 CC testicular, brain, head and neck, mesodermal, gallbladder, rectal,
 CC duodenal, jejunal, ileal, gastric, pancreatic duct, liver duct, gastric
 CC mucosa, gallbladder epithelium, small intestinal mucosa, colorectal
 CC mucosa, pancreatic duct epithelium or liver duct epithelium
 CC preneoplastic/neoplastic disease. AAA16540 to AAA16617 and AAY53228 to
 CC AAY53245 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 459 AA;

Query Match 100.0%; Score 2424; DB 3; Length 459;

Best Local Similarity 100.0%; Pred. No. 1,1e-182; Indels 0; Gaps 0;
Matches 459; Conservative 0; Mismatches 0;

QY 1 MAPLCSPWMLPLIPAPAGLTIVQLLSLLLMFPHQRLPMQDSPLGGSSGSEDDPL 60
DB 1 MAPLCSPWMLPLIPAPAGLTIVQLLSLLLMFPHQRLPMQDSPLGGSSGSEDDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPVEXKSEEGSLKLEDLPTVEAPG 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPVEXKSEEGSLKLEDLPTVEAPG 120
QY 121 DPQEPQNNARHDKGDDQSHMRVGGDPMPRVSPACAGFQSPVDIRQLAFCALAPL 180
DB 121 DPQEPQNNARHDKGDDQSHMRVGGDPMPRVSPACAGFQSPVDIRQLAFCALAPL 180
QY 181 ELIGFQPLPELRLNNGHSVQLTPPGLEMAFGREYRALQHLHMGAAAGRGSEHT 240
DB 181 ELIGFQPLPELRLNNGHSVQLTPPGLEMAFGREYRALQHLHMGAAAGRGSEHT 240
QY 241 VEGHRPFAETHVHLSFAFARVDEALGRPGGLAVLAFLBEGPEENSAYEQLSLBEIA 300
DB 241 VEGHRPFAETHVHLSFAFARVDEALGRPGGLAVLAFLBEGPEENSAYEQLSLBEIA 300
QY 301 EEGSETQVPGDLSALLPDSFSRYFOYEGSLTTPCAQGVITVFNQVMSAKQLHTLS 360
DB 301 EEGSETQVPGDLSALLPDSFSRYFOYEGSLTTPCAQGVITVFNQVMSAKQLHTLS 360
QY 361 DTLMGPDSDRLQNFRAQPLNGRVIEASFPAGVDSPPRAAPVQVNSCLAAGDILALVF 420
DB 361 DTLMGPDSDRLQNFRAQPLNGRVIEASFPAGVDSPPRAAPVQVNSCLAAGDILALVF 420
QY 421 GLFPAVTSVAFLVQMRQRGRGTGKGVSPRAEVAETGA 459
DB 421 GLFPAVTSVAFLVQMRQRGRGTGKGVSPRAEVAETGA 459

RESULT 3

AAB03005
ID AAB03005 standard; protein; 459 AA.

XX AAB03005;

DT 25-SEP-2000 (first entry)

DE Human MN protein.

XX MN protein; tumour associated cell adhesion molecule; oncoprotein;

KM proteoglycan domain; PG domain; carbonic anhydrase; CA domain;

XX abnormal expression; neoplastic disease; cancer; gene therapy.

OS Homo sapiens.

PN WO200024913-A2.

PD 04-MAY-2000.

PF 22-OCT-1999; 99WO-US024879.

PR 23-OCT-1998; 98US-0017776.

XX 23-OCT-1998; 98US-00178115.

PA (FARB) BAYER CORP.

PI (VIR-) INST VIROLOGY.

XX Zavada J, Pastorekova S, Pastorek J;

DR WPI; 2000-350752/30.

XX N-PSDB; AAA52459, AAA52462.

PT A molecule which specifically binds to a site on MN protein (oncoprotein)
PT and prevents adhesion of vertebrate cells to the protein, useful for
PT treating preneoplastic or neoplastic diseases such as cancer.

PS Example 1; Fig 1A-C; 154pp; English.

CC The invention relates to the inhibition of cell adhesion mediated by the
CC MN oncoprotein (also known as the MN/CA IX isoenzyme or the MN/G250
CC protein). The MN protein is a tumour-associated adhesion molecule which
CC comprises a proteoglycan-like (PG) domain (AAB03017) which contains the
CC protein's binding site, and a carbonic anhydrase (CA) domain (AAB03018).
CC Abnormal expression of the MN protein is associated with tumorigenicity.
CC The invention encompasses molecules (e.g., proteins and peptides) which
CC which specifically bind to a site on the MN protein, thereby preventing
CC adhesion of vertebrate cells to the protein in a cell adhesion assay. It
CC also encompasses MN proteins or MN protein fragments which can be added
CC to the extracellular environment to prevent the adhesion of vertebrate
CC cells to each other. The invention also relates to the identification of
CC the binding site of the MN protein and to a method of identifying a site
CC on an MN protein to which cells adhere, comprising testing a series of
CC overlapping peptides from the protein in a cell adhesion assay. The
CC invention encompasses a vector comprising an expression control sequence
CC operatively linked to a nucleic acid encoding the variable domains of a
CC MN-specific antibody, where the domains are separated by a flexible
CC linker peptide (AAB03035) and the vector inhibits the growth of a
CC vertebrate preneoplastic or neoplastic cell that abnormally expresses MN
CC protein. The invention also encompasses a vector comprising a nucleic
CC acid encoding a cytotoxic protein or peptide operatively linked to the MN
CC gene promoter, which inhibits the growth of a vertebrate preneoplastic or
CC neoplastic cell. Also claimed is a repressor complex that binds to the MN
CC gene promoter (AAA52473). MN proteins and peptides, MN-binding proteins
CC and peptides, and expression vectors encoding such proteins and peptides
CC are useful for treating patients with preneoplastic or neoplastic disease
CC (e.g., cancers) associated with or characterised by abnormal MN
CC expression. The present sequence represents the human MN protein.

XX Sequence 459 AA;

Query Match 100.0%; Score 2424; DB 3; Length 459;

Best Local Similarity 100.0%; Pred. No. 1,1e-182; Indels 0; Gaps 0;
Matches 459; Conservative 0; Mismatches 0;

QY 1 MAPLCSPWMLPLIPAPAGLTIVQLLSLLLMFPHQRLPMQDSPLGGSSGSEDDPL 60
DB 1 MAPLCSPWMLPLIPAPAGLTIVQLLSLLLMFPHQRLPMQDSPLGGSSGSEDDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPVEXKSEEGSLKLEDLPTVEAPG 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPVEXKSEEGSLKLEDLPTVEAPG 120
QY 121 DPQEPQNNARHDKGDDQSHMRVGGDPMPRVSPACAGFQSPVDIRQLAFCALAPL 180
DB 121 DPQEPQNNARHDKGDDQSHMRVGGDPMPRVSPACAGFQSPVDIRQLAFCALAPL 180
QY 181 ELIGFQPLPELRLNNGHSVQLTPPGLEMAFGREYRALQHLHMGAAAGRGSEHT 240
DB 181 ELIGFQPLPELRLNNGHSVQLTPPGLEMAFGREYRALQHLHMGAAAGRGSEHT 240
QY 241 VEGHRPFAETHVHLSFAFARVDEALGRPGGLAVLAFLBEGPEENSAYEQLSLBEIA 300
DB 241 VEGHRPFAETHVHLSFAFARVDEALGRPGGLAVLAFLBEGPEENSAYEQLSLBEIA 300
QY 301 EEGSETQVPGDLSALLPDSFSRYFOYEGSLTTPCAQGVITVFNQVMSAKQLHTLS 360
DB 301 EEGSETQVPGDLSALLPDSFSRYFOYEGSLTTPCAQGVITVFNQVMSAKQLHTLS 360
QY 361 DTLMGPDSDRLQNFRAQPLNGRVIEASFPAGVDSPPRAAPVQVNSCLAAGDILALVF 420
DB 361 DTLMGPDSDRLQNFRAQPLNGRVIEASFPAGVDSPPRAAPVQVNSCLAAGDILALVF 420
QY 421 GLFPAVTSVAFLVQMRQRGRGTGKGVSPRAEVAETGA 459
DB 421 GLFPAVTSVAFLVQMRQRGRGTGKGVSPRAEVAETGA 459

RESULT 4

AAE17175

ID AAE17175 standard; protein; 459 AA.
 XX AAE17175;
 AC
 XX
 XX 18-APR-2002 (first entry)
 DT
 XX
 DE Human RCC-associated antigen, G250 protein.
 XX
 KW Human; immune response; immunogenic tumour; renal cell carcinoma;
 KW RCC-associated antigen; G250 protein; immunotherapy; gene therapy;
 KW cancer; cytostatic; immunostimulant.
 XX
 OS Homo sapiens.
 XX
 PN WO200198363-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 20-JUN-2001; 2001MO-NL000461.
 XX
 PR 20-JUN-2000; 2000US-0212669P.
 XX
 PA (UYN1-) UNIV NLMGEGEN.
 XX
 PI Vissers JLM, De Vries JCM, Oosterwijk E, Figdor CG, Adema GJ;
 DR WPI; 2002-139784/18.
 XX
 PT New peptides useful in preparing a composition for the treatment of
 PT cancer, including renal cell carcinoma, cancer of the kidney, prostate,
 PT head, neck, or gastrointestinal tract, and in eliciting immune response.
 XX
 PS Claim 1; Page 35-36; 39pp; English.
 XX
 CC The patent discloses peptides which can be used to elicit an immune
 CC response against a tumour, specifically against an immunogenic tumour.
 CC The peptides are derived from the amino acid sequence of renal cell
 CC carcinoma (RCC)-associated antigen, G250 protein. They are used in
 CC immunotherapy of tumours, in particular renal cell carcinomas. Sequences
 CC of the invention are useful in preparing a composition for the treatment
 CC of cancer, including renal cell carcinoma, cancer of the kidney, head,
 CC prostate, neck, stomach, colon, gastrointestinal tract and bladder. They
 CC are also used for treating immunogenic tumours. Nucleic acid sequences
 CC encoding the peptides of the invention are used in gene therapy. The
 CC present sequence is human RCC-associated antigen, G250 protein
 XX
 SO Sequence 459 AA:
 Query Match 100.0%; Score 2424; DB 5; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1,1e-182;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPPMPLPLIPAPGLTVQLLSLLILMPVAPQGLPMDSDSPYGGSSGSDPL 60
 DB 1 MAPLCSPPMPLPLIPAPGLTVQLLSLLILMPVAPQGLPMDSDSPYGGSSGSDPL 60
 QY 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPOBPONNARHDEKGDQSHMRVGGDPMPWRVSPACAGFQSPVDIRPQLAFCALNPL 180
 DB 121 DPOBPONNARHDEKGDQSHMRVGGDPMPWRVSPACAGFQSPVDIRPQLAFCALNPL 180
 QY 181 ELTGFPQLPPLPELRLNNGHGVQTLTPPGLEVALGGRYRALQLHLHGAAGRGSSHT 240
 DB 181 ELTGFPQLPPLPELRLNNGHGVQTLTPPGLEVALGGRYRALQLHLHGAAGRGSSHT 240
 QY 241 VEGHRFPFAEIHVHLSAFARVDEALGRPGGLAVLAFAEEGPEENSAVEQLSLBEIA 300
 DB 241 VEGHRFPFAEIHVHLSAFARVDEALGRPGGLAVLAFAEEGPEENSAVEQLSLBEIA 300
 QY 301 EEGSETQVPGJLIDISALLPDSRKYFQYEGSLTTPCAQGVITVFNQYMLSAKQLHTLS 360

DB 301 EEGSETQVPGJLIDISALLPDSRKYFQYEGSLTTPCAQGVITVFNQYMLSAKQLHTLS 360
 QY 361 DTLMQPGDSRLQNLNFRATQPLNGRVIEASFPAGVDSFPRAAEVQVNSCLAADILALVF 420
 DB 361 DTLMQPGDSRLQNLNFRATQPLNGRVIEASFPAGVDSFPRAAEVQVNSCLAADILALVF 420
 QY 421 GILFAVTSVAFVLQMRORRGTGKGVSYRPAEVAETGA 459
 DB 421 GILFAVTSVAFVLQMRORRGTGKGVSYRPAEVAETGA 459
 RESULT 5
 ID ABR58596 standard; protein; 459 AA.
 XX ABR58596;
 AC
 XX
 XX 09-JUL-2003 (first entry)
 DT
 XX
 DE Human cancer related protein SEQ ID NO:253.
 XX
 KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
 KW heart disease; atherosclerosis; endometriosis.
 XX
 OS Homo sapiens.
 XX
 PN WO2003025138-A2.
 XX
 PD 27-MAR-2003.
 XX
 PF 17-SEP-2002; 2002WO-US029560.
 XX
 PR 17-SEP-2001; 2001US-0323469P.
 PR 20-SEP-2001; 2001US-0323887P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 08-FEB-2002; 2002US-0358145P.
 PR 08-FEB-2002; 2002US-0355257P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (BOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aiz N, Glsh KC, Hevezl PA, Mack DH, Wilson KE;
 PI Zlotnik A;
 XX
 XX WPI; 2003-354600/33.
 DR N-PSDB; ACC72730.
 XX
 PT New genes that are up-regulated or down-regulated in cancers, useful as
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 PT therapeutic targets for screening drugs for treating these diseases.
 XX
 PS Claim 12; Page 744; 767pp; English.
 XX
 CC The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
 CC related gene nucleotide sequences which encode the proteins given in
 CC ABR58596. Also described: (1) determining the presence or
 CC absence of a pathological cell in a patient; (2) an expression vector
 CC comprising a nucleic acid molecule described above; (3) a host cell
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
 CC of (4); (6) specifically targeting a compound to a pathological cell in a
 CC patient by administering to the patient the antibody above; and (7) a
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or
 CC therapeutic targets. In particular, the nucleic acid is useful for
 CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
 CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
 CC atherosclerosis and endometriosis. The nucleic acid is also useful in
 CC drug screening, particularly for identifying agents for treating these

CC pathologies
XX Sequence 459 AA;
SQ

Query Match 100.0%; Score 2424; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 1,1e-182;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPICSPWMLPLIPAPAPGLTVQLLSLILMPVHQRPLPMQBDSPICGSSSGEDDL 60
DB 1 MAPICSPWMLPLIPAPAPGLTVQLLSLILMPVHQRPLPMQBDSPICGSSSGEDDL 60
QY 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPVKKSEEGSLKLEDLPTVEARG 120
DB 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPVKKSEEGSLKLEDLPTVEARG 120
QY 121 DPOEPQNNARHDKGDDQSHWRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPOEPQNNARHDKGDDQSHWRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELIGFOLPLPELRLNNGHSVQLTPPGLEMALPGREYRALQLHMGAGRPGSEHT 240
DB 181 ELIGFOLPLPELRLNNGHSVQLTPPGLEMALPGREYRALQLHMGAGRPGSEHT 240
QY 241 VEGHFRPAEIHVHLSLTPARVDALGRPGGLAVLAFLKEGPEENSAYEQLSRLLEEA 300
DB 241 VEGHFRPAEIHVHLSLTPARVDALGRPGGLAVLAFLKEGPEENSAYEQLSRLLEEA 300
QY 301 EEGSETQVPGDLISALLPSDFSRFYQYEGSLTPPCAQVIMTVFNQVWLSAKOHLTIS 360
DB 301 EEGSETQVPGDLISALLPSDFSRFYQYEGSLTPPCAQVIMTVFNQVWLSAKOHLTIS 360
QY 361 DTLWGPDSRLQLNFRATQPLNGRVIEASPPAGVDSPPRAAPVQVNSCLAAGDILALVF 420
DB 361 DTLWGPDSRLQLNFRATQPLNGRVIEASPPAGVDSPPRAAPVQVNSCLAAGDILALVF 420
QY 421 GLFPAVTSVAFVQMRQHRGRTKGVSYPRAEVAETGA 459
DB 421 GLFPAVTSVAFVQMRQHRGRTKGVSYPRAEVAETGA 459

RESULT 6
ABP97744
ID ABP97744 standard; protein; 459 AA.
XX
AC ABP97744;
XX
DT 28-MAY-2003 (first entry)
XX
DE Amino acid sequence of human CA9 polypeptide.
XX
KM Human; angiogenesis; wound healing; retinopathy; ischemia; inflammation;
KM microvasculopathy; bone healing; skin inflammation; HOG3; HOG8; HOG16;
KM follicular development; CA9; HXB; IGFBP5; HFRAP; STC1; mlg-6; SSR4;
KM cancer.
XX
OS Homo sapiens.
XX
PN MO2003010205-AL.
PD 06-FEB-2003.
XX
PF 26-JUL-2002; 2002MO-US023786.
XX
PR 26-JUL-2001; 2001US-0307600P.
PR 24-JUL-2002; 2002US-00201642.
XX
PA (UNIV-) UNIV DUKE MEDICAL CENT.
XX
PI Riggins GJ, Lai A;
XX
DR WPI; 2003-239423/23.
DR N-PSDB; AB277284.

XX Inhibiting angiogenesis for treating wound healing, retinopathy,
XX ischtema, inflammation, microvasculopathy, bone healing, skin
XX inflammation or follicular development by providing to a subject an
XX antisense polynucleotide.
XX
PS Claim 4; Page 47-48; 66pp; English.
XX
The present sequence is a human CA9 polypeptide. It is used in the method
CC of the invention. The specification describes a method modulating
CC angiogenesis associated with wound healing, retinopathy, ischemia,
CC inflammation, microvasculopathy, bone healing, skin inflammation or
CC follicular development. The method comprises providing to a subject HOG3,
CC HOG8, HOG18, CA9, HXB, IGFBP5, HFRAP, STC1, mlg-6 or SSR4. The methods,
CC antisense polynucleotides, polypeptides and antibodies are useful for
CC treating wound healing, retinopathy, ischemia, inflammation,
CC microvasculopathy, bone healing, skin inflammation or follicular
CC development, or cancer such as breast, colon or lung cancer, or
CC glioblastoma
XX
SQ Sequence 459 AA;
Query Match 100.0%; Score 2424; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 1,1e-182;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPICSPWMLPLIPAPAPGLTVQLLSLILMPVHQRPLPMQBDSPICGSSSGEDDL 60
DB 1 MAPICSPWMLPLIPAPAPGLTVQLLSLILMPVHQRPLPMQBDSPICGSSSGEDDL 60
QY 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPVKKSEEGSLKLEDLPTVEARG 120
DB 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPVKKSEEGSLKLEDLPTVEARG 120
QY 121 DPOEPQNNARHDKGDDQSHWRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPOEPQNNARHDKGDDQSHWRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELIGFOLPLPELRLNNGHSVQLTPPGLEMALPGREYRALQLHMGAGRPGSEHT 240
DB 181 ELIGFOLPLPELRLNNGHSVQLTPPGLEMALPGREYRALQLHMGAGRPGSEHT 240
QY 241 VEGHFRPAEIHVHLSLTPARVDALGRPGGLAVLAFLKEGPEENSAYEQLSRLLEEA 300
DB 241 VEGHFRPAEIHVHLSLTPARVDALGRPGGLAVLAFLKEGPEENSAYEQLSRLLEEA 300
QY 301 EEGSETQVPGDLISALLPSDFSRFYQYEGSLTPPCAQVIMTVFNQVWLSAKOHLTIS 360
DB 301 EEGSETQVPGDLISALLPSDFSRFYQYEGSLTPPCAQVIMTVFNQVWLSAKOHLTIS 360
QY 361 DTLWGPDSRLQLNFRATQPLNGRVIEASPPAGVDSPPRAAPVQVNSCLAAGDILALVF 420
DB 361 DTLWGPDSRLQLNFRATQPLNGRVIEASPPAGVDSPPRAAPVQVNSCLAAGDILALVF 420
QY 421 GLFPAVTSVAFVQMRQHRGRTKGVSYPRAEVAETGA 459
DB 421 GLFPAVTSVAFVQMRQHRGRTKGVSYPRAEVAETGA 459

RESULT 7
ABUS6656
ID ABUS6656 standard; protein; 459 AA.
XX
AC ABUS6656;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polypeptide #249.
XX
KM Lung cancer-associated polypeptide; cytostatic; emphysema;
KM antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KM small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KM chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 OS
 XX WO200286443-A2.
 EN
 XX 31-OCT-2002.
 PD
 XX 18-APR-2002; 2002WO-US012476.
 PF
 XX 18-APR-2001; 2001US-0284770P.
 PR
 XX 10-MAY-2001; 2001US-0290492P.
 PR
 XX 09-NOV-2001; 2001US-0339245P.
 PR
 XX 13-NOV-2001; 2001US-0350666P.
 PR
 XX 29-NOV-2001; 2001US-034370P.
 PR
 XX 12-APR-2002; 2002US-0372246P.
 PA
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PI
 XX Aziz N, Murray R;
 PI
 XX WPI; 2003-093161/08.
 DR
 XX N-PSDB; ABX76385.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 PT
 XX
 PS Claim 27; Page 380; 453pp; English.
 XX
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences of
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
 CC invention
 CC
 XX
 SQ Sequence 459 AA;
 Query Match 100.0%; Score 2424; DB 6; Length 459;
 Best local similarity 100.0%; Pred. No. 1, 1e-182;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAPLCPSFWLPLIPAPAPGLTVQLLSTLLMPVPOQLPFWQSDSPGLGSSGSDPPL 60
 DB 1 MAPLCPSFWLPLIPAPAPGLTVQLLSTLLMPVPOQLPFWQSDSPGLGSSGSDPPL 60
 QY 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPOBPNNAHDKKGDQSHMRXGDPWPWRVSPACAGSFQSPVDIRPOLAFCALRPL 180
 DB 121 DPOBPNNAHDKKGDQSHMRXGDPWPWRVSPACAGSFQSPVDIRPOLAFCALRPL 180
 QY 181 ELIGFOPLPLBELRLRNHGSVQTLPLGLSMALGGRERYRALQHLHMGAGRGSGSHT 240
 DB 181 ELIGFOPLPLBELRLRNHGSVQTLPLGLSMALGGRERYRALQHLHMGAGRGSGSHT 240

QY 241 VEGHREPAIHVHLSTAFARVDEALGRPGGLAVLAAPLEGEPPENSAYEQLLSRLEEA 300
 DB 241 VEGHREPAIHVHLSTAFARVDEALGRPGGLAVLAAPLEGEPPENSAYEQLLSRLEEA 300
 QY 301 EBSGETQVPGDLISALPBDPSRRYQYFGSLTTPPCAGQVITVYNQTMISAKQHTLS 360
 DB 301 EBSGETQVPGDLISALPBDPSRRYQYFGSLTTPPCAGQVITVYNQTMISAKQHTLS 360
 QY 361 DTLMGPGDSRLQINRATOPUNGRVTEASFPAGVDSFPRAEPQUNSCLAAGGILALVF 420
 DB 361 DTLMGPGDSRLQINRATOPUNGRVTEASFPAGVDSFPRAEPQUNSCLAAGGILALVF 420
 QY 421 GLIFAVTSVAFVQMRQRRGTGCGVSYPRAEVAETGA 459
 DB 421 GLIFAVTSVAFVQMRQRRGTGCGVSYPRAEVAETGA 459
 RESULT 8
 ID ABUS6408 standard; protein; 459 AA.
 XX
 AC ABUS6408;
 XX
 DT 02-APR-2003 (first entry)
 DE Lung cancer-associated polypeptide #1.
 XX
 XX Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiaesthetic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 OS
 XX Unidentified.
 OS
 XX WO200286443-A2.
 EN
 XX 31-OCT-2002.
 PD
 XX 18-APR-2002; 2002WO-US012476.
 PF
 XX 18-APR-2001; 2001US-0284770P.
 PR
 XX 10-MAY-2001; 2001US-0290492P.
 PR
 XX 09-NOV-2001; 2001US-0339245P.
 PR
 XX 13-NOV-2001; 2001US-0350666P.
 PR
 XX 29-NOV-2001; 2001US-034370P.
 PR
 XX 12-APR-2002; 2002US-0372246P.
 PA
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PI
 XX Aziz N, Murray R;
 PI
 XX WPI; 2003-093161/08.
 DR
 XX N-PSDB; ABX76124.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 PT
 XX
 PS Claim 27; Page 189; 453pp; English.
 XX
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell

CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, actinia and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
CC invention
CC
XX
XX
SQ Sequence 459 AA;

Query Match 100.0%; Score 2424; DB 6; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.1e-182;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLLPAPAPGLTVQLLSLLMLPVHFORLPRMOEDSPFGGSSGSDPDL 60
DB 1 MAPLCSPWMLPLLPAPAPGLTVQLLSLLMLPVHFORLPRMOEDSPFGGSSGSDPDL 60

QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DPEPQNNARHDKEDGDSHMRVYGGDPMPRVSPACAGRQSPVDIRPOLAFCALRPL 180
DB 121 DPEPQNNARHDKEDGDSHMRVYGGDPMPRVSPACAGRQSPVDIRPOLAFCALRPL 180

QY 181 ELIGFOLPPLPELRLNNGHVSQVLTLPGLLENALGGRREYRALQHLHGAAGRGSEHT 240
DB 181 ELIGFOLPPLPELRLNNGHVSQVLTLPGLLENALGGRREYRALQHLHGAAGRGSEHT 240

QY 241 VEGHRFPAEIHVHVLSTAFARVDEALGRPGGLAVLAFLKEGSEENSAYEQLLSLREIA 300
DB 241 VEGHRFPAEIHVHVLSTAFARVDEALGRPGGLAVLAFLKEGSEENSAYEQLLSLREIA 300

QY 301 EEGSETOVPGGLDISALLPDSFSRYFYEGSLTTPCAQGVITVFNQVMSAKQLHTLS 360
DB 301 EEGSETOVPGGLDISALLPDSFSRYFYEGSLTTPCAQGVITVFNQVMSAKQLHTLS 360

QY 361 DTLWPGDSRLQLNFRATQPLNGRVLEASFPAVDSPRAAEVQVQNSCLAAGDILAVF 420
DB 361 DTLWPGDSRLQLNFRATQPLNGRVLEASFPAVDSPRAAEVQVQNSCLAAGDILAVF 420

QY 421 GLIFAVTSVAFLVOMRQRGRGTGVSYPAEVATGA 459
DB 421 GLIFAVTSVAFLVOMRQRGRGTGVSYPAEVATGA 459

RESULT 9
ID AAB82848 standard; protein; 610 AA.
XX
XX AAB82848;
AC
XX
XX 12-NOV-2001 (first entry)
DE
XX Kidney cancer specific antigen G250-GM-CSF fusion protein.
XX
XX GM-CSF, granulocyte-macrophage colony stimulating factor; G250; vaccine;
KW genetic immunisation; adoptive immunotherapy; therapy; kidney cancer;
KW antigen; renal cell carcinoma; tumour; human; antitumour.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX
XX Key Location/Qualifiers
FT 1..144
FT /label= G250
FT /label= 145..146
FT /label= Linker
FT Misc-difference 145
FT /note= "encoded by GCG"
FT Misc-difference 146

FT /note= "encoded by GCG"
FT Protein 147..604
FT /label= GM-CSF
FT Peptide 605..610
FT /note= "histidine tag, not encoded by the nucleic acid of
FT AAB82848"
XX
XX
XX WO200160317-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004595.
XX
XX 14-FEB-2000; 2000US-0182429P.
XX 15-FEB-2000; 2000US-0182636P.
XX
XX (RBC) UNITV CALIFORNIA.
XX
XX Belldegrin A, Tso C;
XX
XX WPI; 2001-557625/62.
XX N-PSDB; AAB26551.
XX
XX
XX New construct with a G250 kidney cancer specific antigen attached to a
XX granulocyte macrophage colony stimulating factor, useful in gene therapy
XX or as a vaccine for treating renal cell cancers.
XX
XX Claim 9; Page 75; 78pp; English.
XX
XX The present sequence is that of a fusion protein composed of the G250
XX kidney cancer specific antigen attached via a dipeptide linker to human
XX granulocyte macrophage colony stimulating factor (GM-CSF), plus a C-
XX terminal hexahistidine tag. The fusion protein can be obtained by
XX recombinant expression, e.g. in insect Sf8 cells, using a fusion
XX construct (see AAB26551) obtained from human G250 and GM-CSF cDNAs. The
XX fusion protein can be used to raise an immune response directed against
XX renal cell cancers. It is used as a vaccine or in adoptive immunotherapy.
XX Nucleic acids encoding the fusion protein can be used as DNA vaccines, to
XX transfect cells in an adoptive immunotherapy treatment regimen, and in
XX gene therapy for treating or preventing renal cell cancers
XX
XX
SQ Sequence 610 AA;

Query Match 99.8%; Score 2419; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 4e-182;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 APLCSPWMLPLLPAPAPGLTVQLLSLLMLPVHFORLPRMOEDSPFGGSSGSDPDL 61
DB 147 APLCSPWMLPLLPAPAPGLTVQLLSLLMLPVHFORLPRMOEDSPFGGSSGSDPDL 61

QY 62 EEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 121
DB 62 EEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 121

QY 122 DPEPQNNARHDKEDGDSHMRVYGGDPMPRVSPACAGRQSPVDIRPOLAFCALRPL 181
DB 122 DPEPQNNARHDKEDGDSHMRVYGGDPMPRVSPACAGRQSPVDIRPOLAFCALRPL 181

QY 182 LIGFOLPPLPELRLNNGHVSQVLTLPGLLENALGGRREYRALQHLHGAAGRGSEHT 241
DB 182 LIGFOLPPLPELRLNNGHVSQVLTLPGLLENALGGRREYRALQHLHGAAGRGSEHT 241

QY 242 EGHRRFPAEIHVHVLSTAFARVDEALGRPGGLAVLAFLKEGSEENSAYEQLLSLREIA 301
DB 242 EGHRRFPAEIHVHVLSTAFARVDEALGRPGGLAVLAFLKEGSEENSAYEQLLSLREIA 301

QY 302 EEGSETOVPGGLDISALLPDSFSRYFYEGSLTTPCAQGVITVFNQVMSAKQLHTLS 361
DB 302 EEGSETOVPGGLDISALLPDSFSRYFYEGSLTTPCAQGVITVFNQVMSAKQLHTLS 361

QY 362 DTLWPGDSRLQLNFRATQPLNGRVLEASFPAVDSPRAAEVQVQNSCLAAGDILAVF 421
DB 362 DTLWPGDSRLQLNFRATQPLNGRVLEASFPAVDSPRAAEVQVQNSCLAAGDILAVF 421

Db 507 TLWPGDSRLQNLFRATQPLNGRVIKASPPAGVDSPPRAAEVQNSCLAADILATVFG 566

QY 422 LLFAVTSVAFLVQMRQRHGRGTGKGVSYRAAEVAETGA 459

Db 567 LLFAVTSVAFLVQMRQRHGRGTGKGVSYRAAEVAETGA 604

RESULT 10
AAV53245
ID AAV53245 standard; protein: 377 AA.

AC AAV53245;

XX 16-JUN-2000 (first entry)

DE MN protein extracellular domain SEQ ID NO:87.

XX Human; MN protein; MN gene; oncogene; carbonic anhydrase; tumour;
KM oncogenesis; diagnosis; neoplastic disease; cancer; carcinoma;
KM MN/CA IX isoenzyme.

XX Homo sapiens.

OS US6027887-A.

XX 22-FEB-2000.

PD 24-JAN-1997; 97US-00787739.

XX 21-OCT-1992; 92US-00964589.
PR 30-DEC-1993; 93US-00177093.
PR 15-JUN-1994; 94US-00260190.
PR 07-JUN-1995; 95US-00477504.
PR 07-JUN-1995; 95US-00481658.
PR 07-JUN-1995; 95US-00485049.
PR 07-JUN-1995; 95US-00485862.
PR 07-JUN-1995; 95US-00485863.
PR 07-JUN-1995; 95US-00486756.
PR 07-JUN-1995; 95US-00487077.

XX (SLSC-) SLOVAK ACAD SCI INST VIROLOGY.

XX Pastorek J, Zavada J, Pastorekova S;

XX WPI; 2000-194827/17.

XX Nucleic acid based assay for diagnosing a wide variety of
PT preneoplastic/neoplastic disease comprises screening for the presence of
PT abnormal MN gene expression in a vertebrate.

XX Disclosure; Col 16; 87pp; English.

XX The present invention describes a method of screening for
CC preneoplastic/neoplastic disease. The method comprises: (1) determining
CC whether abnormal MN gene expression is present in a vertebrate; and (2)
CC if abnormal MN gene expression is determined to be present in the
CC vertebrate, determining that the vertebrate has a significant risk of
CC having preneoplastic/neoplastic disease. The MN gene is an oncogene and
CC encodes an MN protein (also referred to as MN/CA IX isoenzyme). The MN
CC protein is a tumour associated carbonic anhydrase isoenzyme. The method
CC is used for detecting a wide variety of preneoplastic/neoplastic diseases
CC in a vertebrate, preferably a human. The disease detected is mammary,
CC bladder, renal, urinary tract, ovarian, uterine, cervical, endometrial,
CC vaginal, vulval, prostate, liver, lung, skin, thyroid, pancreatic,
CC testicular, brain, head and neck, mesodermal, gallbladder, rectal,
CC duodenal, jejunal, ileal, gastric, pancreatic duct, liver duct, gastric
CC mucosa, gallbladder epithelium, small intestinal mucosa, colorectal
CC mucosa, pancreatic duct epithelium or liver duct epithelium
CC preneoplastic/neoplastic disease. AAV5340 to AAV561 and AAV53228 to
CC AAV53245 represent sequences used in the exemplification of the present
XX invention
XX Sequence 377 AA;

Query Match 82.8%; Score 2007; DB 3; Length 377;
Best Local Similarity 100.0%; Pred. No. 6,9e-150;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 ORLPRMQEDSPDGSGSDGDDPLGEBDLPSEDSPPREDDPGEBDLPGEEDLPGEEDLP 97

Db 1 QRLPRMQEDSPDGSGSGEDDPLGEBDLPSEDSPPREDDPGEBDLPGEEDLPGEEDLP 60

QY 98 VPKRSEEGSLLEBULPYVEAPGDPQEPQNNANHRDKEDDSSHRYGGDPPWPVSPACA 157

Db 61 VPKRSEEGSLLEBULPYVEAPGDPQEPQNNANHRDKEDDSSHRYGGDPPWPVSPACA 120

QY 158 GRFQSPVDIRPOLAFCALRPLELGFQLPPLPELRNRNNGHVQTLPPGLEMALGPG 217

Db 121 GRFQSPVDIRPOLAFCALRPLELGFQLPPLPELRNRNNGHVQTLPPGLEMALGPG 180

QY 218 REYRALQHLHWGAAGRQSGHTEGHRFPPELIVVHISTFAVDEALGRPGGLAVLAA 277

Db 181 REYRALQHLHWGAAGRQSGHTEGHRFPPELIVVHISTFAVDEALGRPGGLAVLAA 240

QY 278 FLEBGPENSAYEQLLSRLEIAEGSETOVPGLDISALPSPRSRYFOYEGSLTTPPCA 337

Db 241 FLEBGPENSAYEQLLSRLEIAEGSETOVPGLDISALPSPRSRYFOYEGSLTTPPCA 300

QY 338 QGVIVTFNQVTMLSAKQLHTLSDTLWPGDSRLQNLFRATQPLNGRVIKASPPAGVDS 397

Db 301 QGVIVTFNQVTMLSAKQLHTLSDTLWPGDSRLQNLFRATQPLNGRVIKASPPAGVDS 360

QY 398 PRAAEVQNSCLAAGD 414

Db 361 PRAAEVQNSCLAAGD 377

RESULT 11
ID AAB03021
AAV53021 standard; protein: 377 AA.

XX AAB03021;

XX 25-SEP-2000 (first entry)

XX Mature human MN protein N-terminal extracellular domain.

XX MN protein; tumour associated cell adhesion molecule; oncoprotein;
KM proteoglycan domain; PG domain; carbonic anhydrase; CA domain;
KM abnormal expression; neoplastic disease; cancer; gene therapy.

XX Homo sapiens.

XX WO200024913-A2.

XX 04-MAY-2000.

XX 22-OCT-1999; 99WO-US024879.

XX 23-OCT-1998; 98US-00177776.
PR 23-OCT-1998; 98US-00178115.

XX (FAPB) BAYER CORP.
PA (VIRO-) INST VIROLOGY.

XX Zavada J, Pastorekova S, Pastorek J;

XX WPI; 2000-150752/30.

XX A molecule which specifically binds to a site on MN protein (oncoprotein)
PT and prevents adhesion of vertebrate cells to the protein, useful for
PT creating preneoplastic or neoplastic diseases such as cancer.

XX Disclosure; Page 137-138; 154pp; English.

XX The invention relates to the inhibition of cell adhesion mediated by the

CC MN oncoprotein (also known as the MN/CA IX isoenzyme or the MN/G250
 CC protein). The MN protein is a tumour-associated adhesion molecule which
 CC comprises a proteoglycan-like (PG) domain (AAB03017) which contains the
 CC protein's binding site, and a carbonic anhydrase (CA) domain (AAB03018).
 CC Abnormal expression of the MN protein is associated with tumorigenicity.
 CC The invention encompasses molecules (e.g., proteins and peptides) which
 CC which specifically bind to a site on the MN protein, thereby preventing
 CC adhesion of vertebrate cells to the protein in a cell adhesion assay. It
 CC also encompasses MN proteins or MN protein fragments which can be added
 CC to the extracellular environment to prevent the adhesion of vertebrate
 CC cells to each other. The invention also relates to the identification of
 CC the binding site of the MN protein and to a method of identifying a site
 CC on an MN protein to which cells adhere, comprising testing a series of
 CC overlapping peptides from the protein in a cell adhesion assay. The
 CC invention encompasses a vector comprising an expression control sequence
 CC operatively linked to a nucleic acid encoding the variable domains of a
 CC MN-specific antibody, where the domains are separated by a flexible
 CC linker peptide (AAB03035) and the vector inhibits the growth of a
 CC vertebrate preneoplastic or neoplastic cell that abnormally expresses MN
 CC protein. The invention also encompasses a vector comprising a nucleic
 CC acid encoding a cytotoxic protein or peptide operatively linked to the MN
 CC gene promoter, which inhibits the growth of a vertebrate preneoplastic or
 CC neoplastic cell. Also claimed is a repressor complex that binds to the MN
 CC gene promoter (AAB02473). MN proteins and peptides, MN-binding proteins
 CC and peptides, and expression vectors encoding such proteins and peptides
 CC are useful for treating patients with preneoplastic or neoplastic disease
 CC (e.g., cancers) associated with or characterized by abnormal MN
 CC expression. Sequences AAB03006, AAB03015-B03022, AAB03022-B03031,
 CC AAB03036 and AAB03059-B03060 represent specific domains or regions of the
 CC human MN protein (AAB03005) referred to in the invention

XX Sequence 377 AA;

Query Match 82.8%; Score 2007; DB 3; Length 377;
 Best Local Similarity 100.0%; Pred. No. 6.9e-150;
 Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 QRLPRMOEDSPLOGSSGSGEDDPLGEEDLPSEEDSPREEDLPGEEDLPGEEDLP 97
 DB 1 QRLPRMOEDSPLOGSSGSGEDDPLGEEDLPSEEDSPREEDLPGEEDLPGEEDLP 60
 QY 98 VKRSEBEGSLKIEDLPTVEAPDPOEPONNARXKGGDQSHWYGGDPPMPRSPACA 157
 DB 61 VKRSEBEGSLKIEDLPTVEAPDPOEPONNARXKGGDQSHWYGGDPPMPRSPACA 120
 QY 158 GRFQSPVDIRPOLAFCPALRPTELIGFQLPPLPELRLRNNGHSVQTLTPGLEMAAGPG 217
 DB 121 GRFQSPVDIRPOLAFCPALRPTELIGFQLPPLPELRLRNNGHSVQTLTPGLEMAAGPG 180
 QY 218 REVRALQHLHMGANARPGSEHTEVEGRFPALHYVHLSTARVDEALGRPGGLAVTAA 277
 DB 181 REVRALQHLHMGANARPGSEHTEVEGRFPALHYVHLSTARVDEALGRPGGLAVTAA 240
 QY 278 FLEEGPENSAVYQLLSLEELAEESGSETOVFGDLISALLPSDFSRFYQYGSGLTPPCA 337
 DB 241 FLEEGPENSAVYQLLSLEELAEESGSETOVFGDLISALLPSDFSRFYQYGSGLTPPCA 300
 QY 338 QGVIMTVENOTVWLSKQHLTSDTLMGPGDSRLQNPANQPLNGRVTASFPAGVDS 397
 DB 301 QGVIMTVENOTVWLSKQHLTSDTLMGPGDSRLQNPANQPLNGRVTASFPAGVDS 360
 QY 398 PRAAEPYQLNSCLAAGD 414
 DB 361 PRAAEPYQLNSCLAAGD 377

RESULT 12
 AAR41746 standard; protein; 429 AA.
 ID AAR41746;
 AC AAR41746;
 XX
 XX 25-MAR-2003 (revised)

DT 25-MAR-1994 (first entry)
 XX
 DE MN protein.
 XX
 KM MN; endogenous; Matu; quasi-viral agent; human; mammary tumour; prion;
 KM classical virus; slow virus; exogenous MX; p58X; cytoplasmic antigen;
 KM conservative; Hela cell; twin protein; p54/58N; cell surface; nucleus;
 KM monoclonal antibody; Mab M75; neoplasm; pre-neoplastic disease; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN W09318152-A1.
 XX
 PD 16-SEP-1993.
 XX
 PF 08-MAR-1993; 93WO-US002024.
 XX
 PR 11-MAR-1992; 92CS-00000709.
 PR 21-OCT-1992; 92US-00964589.
 XX
 PA (CIBA) CIBA CORNING DIAGNOSTICS CORP.
 PA (VIRO-) INST VIROLOGY.
 XX
 PT Zavada J, Pastorekova S, Pastorek J;
 DR MPI; 1993-303466/38.
 XX
 DR N-PSDB; AAQ8456.
 XX
 PT New MN gene and polypeptide(s) - used in diagnosis, prognosis and therapy
 PT of neoplastic and/or pre-neoplastic disease.
 PS
 PS Claim 7; Fig 1; 72pp; English.

CC This sequence is encoded by the intronless MN gene which is a cellular
 CC gene which is the endogenous component of the Matu agent. Matu is a novel
 CC quasi-viral agent with rather unusual properties. It is presumably
 CC derived from a human mammary tumour. In some aspects it resembles
 CC classical viruses, whereas in other respects it resembles "slow" viruses
 CC (prions), and in still other aspects it is different from both classes of
 CC viruses. Matu is a two component system. One part of the complex,
 CC exogenous MX, is transmissible, and is manifest by a protein, p58X, which
 CC is a cytoplasmic antigen which reacts with some natural sera, of humans
 CC and of various animals. The other component, MN, is endogenous to human
 CC cells. MN is a cellular gene showing very little homology with known DNA
 CC sequences. It is rather conservative and present as a single copy in the
 CC chromosomal DNA of various vertebrates. MN is manifest in Hela cells by a
 CC twin protein p54/58N, that is localised on the cell surface and in the
 CC nucleus. Immunoblots using a monoclonal antibody reactive with p54/58N
 CC (Mab M75) reveals two bands at 54 kD and 58 kD. These two bands may
 CC correspond to one type of protein that differs by glycosylation pattern
 CC or by how it is processed. The expression of the MN gene is strongly
 CC correlated with tumorigenicity. MN products can be used in can be used
 CC in diagnostic and/or prognostic assays for neoplastic and/or pre-
 CC neoplastic disease. MN polypeptides, produced recombinantly by
 CC unicellular hosts, can also be used for antibody production and in
 CC vaccines for inducing protective immunity against neoplastic disease and
 CC a dampening effect upon tumorigenic activity. (Updated on 25-MAR-2003 to
 CC correct PN field.)

XX Sequence 429 AA;
 SQ
 Query Match 57.7%; Score 1398.5; DB 2; Length 429;
 Best Local Similarity 69.1%; Pred. No. 8.7e-102;
 Matches 300; Conservative 12; Mismatches 83; Indels 39; Gaps 7;

QY 38 QRLPRMOEDSPLOGSSGSGEDDPLGEEDLPSEEDSPREEDLPGEEDLPGEEDLP 96
 DB 1 QRLPRMOEDSPLEALLGKMTWARRICVRRIRH-ERRITLERITVIRITL 59
 QY 97 EVKPSBEGSLKIEDLPTVEAPDPOEPONNARXKGGDQSHWYGGDPPMPW----- 150
 DB 60 KIMPSBEGSLKIEDLPTVEAPDPOEPONNARXKGGDQSHWYGGDPPGCPQPA 119

QY 151 --RVSPACAGRFQSPVDIRPOLAFC-----PALRPLELLGFQPLPELRLRNNGH 200
 Db 120 RAASSPRWMSAPSSPSARP-----CAPMNSWASSSRSSQNCACQMP----- 162
 QY 201 SVQTLTPGLEMLAGGR--YRALQLHMGAGAPGSGEHTVEGHRFPAAIHVHSTA 256
 Db 163 --CCATDPASWARDGSGSRAGVAPALQLHMGAGAPGSGEHTVEGHRFPAAIHVHSTA 220
 QY 259 FAVDEALGPGGLAVLAFALEEGPENSAVEQLLSRLBEIAEGSETOVGLISALP 318
 Db 221 FAVDEALGPGGLAVLAFAFMARAKKTQSYQLLSRLBEIAEGSETOVGLISALP 280
 QY 319 SDESRFYQYEGSLITTPCAQGVITVFNQVWMLSAKQHLTSDTLWGDSRLQINFRAT 378
 Db 281 SDESRFYQYEGSLITTPCAQGVITVFNQVWMLSAKQHLTSDTLWGDSRLQINFRAT 340
 QY 379 QPLNGRVIEASFPAGVDSPPRAAEFYQVNSCLAADITLAVFGLEFNTSVAFVQMRQ 438
 Db 341 QPLNGRVIEASFPAGVDSPPRAAEFYQVNSCLAADITLAVFGLEFNTSVAFVQMRQ 400
 QY 439 HRRGRTKGVSYPRA 452
 Db 401 HRRGRTKGVSYPSPS 414

RESULT 13

AA53241

ID AA53241 standard; protein; 257 AA.

AC AA53241;

DT 16-JUN-2000 (first entry)

DE MN protein carbonic anhydrase domain SEQ ID NO:51.

KM Human; MN protein; MN gene; oncogene; carbonic anhydrase; tumour;
 KW oncogenesis; diagnosis; neoplastic disease; cancer; carcinoma;
 XX MN/CA IX isoenzyme.

OS Homo sapiens.

PN US6027867-A.

PD 22-FEB-2000.

PF 24-JAN-1997; 97US-00787739.

PR 21-OCT-1992; 92US-00964589.

PR 30-DEC-1993; 93US-00177093.

PR 15-JUN-1994; 94US-00260190.

PR 07-JUN-1995; 95US-00477504.

PR 07-JUN-1995; 95US-00481658.

PR 07-JUN-1995; 95US-00485049.

PR 07-JUN-1995; 95US-00485862.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00485863.

CC 1: abnormal MN gene expression is determined to be present in the
 CC vertebrate, determining that the vertebrate has a significant risk of
 CC having preneoplastic/neoplastic disease. The MN gene is an oncogene and
 CC encodes an MN protein (also referred to as MN/CA IX isoenzyme). The MN
 CC protein is a tumor associated carbonic anhydrase isoenzyme. The method
 CC is used for detecting a wide variety of preneoplastic/neoplastic diseases
 CC in a vertebrate, preferably a human. The disease detected is mammary,
 CC bladder, renal, urinary tract, ovarian, uterine, cervical, endometrial,
 CC vaginal, vulval, prostate, liver, lung, skin, thyroid, pancreatic,
 CC testicular, brain, head and neck, mesodermal, gallbladder, rectal,
 CC duodenal, jejunal, ileal, gastric, pancreatic duct, liver duct, gastric
 CC mucosa, gallbladder epithelium, small intestinal mucosa, colorectal
 CC mucosa, pancreatic duct epithelium or liver duct epithelium
 CC preneoplastic/neoplastic disease. AA16540 to AA16617 and AA53228 to
 CC AA53245 represent sequences used in the exemplification of the present
 CC invention
 SQ Sequence 257 AA;

Query Match 56.5%; Score 1370; DB 3; Length 257;
 Best Local Similarity 100.0%; Pred No. 7.8e-100; Indels 0; Gaps 0;
 Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 GDDSHMRVGGDPMPVSPACAGRFQSPVDIRPOLAFCPALRPLELLGFQPLPELR 194
 Db 1 GDDSHMRVGGDPMPVSPACAGRFQSPVDIRPOLAFCPALRPLELLGFQPLPELR 60
 QY 195 LRNGHGVQTLTPGLEMLAGPREYRALQLHMGAGAPGSGEHTVEGHRFPAAIHVH 254
 Db 61 LRNGHGVQTLTPGLEMLAGPREYRALQLHMGAGAPGSGEHTVEGHRFPAAIHVH 120
 QY 255 LSTAFARVDEALGPGGLAVLAFALEEGPENSAVEQLLSRLBEIAEGSETOVGLDIS 314
 Db 121 LSTAFARVDEALGPGGLAVLAFALEEGPENSAVEQLLSRLBEIAEGSETOVGLDIS 180
 QY 315 ALTPSDSRFYQYEGSLITTPCAQGVITVFNQVWMLSAKQHLTSDTLWGDSRLQIN 374
 Db 181 ALTPSDSRFYQYEGSLITTPCAQGVITVFNQVWMLSAKQHLTSDTLWGDSRLQIN 240
 QY 375 FRATQPLNGRVIEASFP 391
 Db 241 FRATQPLNGRVIEASFP 257

RESULT 14

ID AAB03018 standard; protein; 257 AA.

AC AAB03018;

DT 25-SEP-2000 (first entry)

DE Human MN protein carbonic anhydrase (CA) domain.

KM MN protein; tumour associated cell adhesion molecule; oncoprotein;
 KW proteoglycan domain; PG domain; carbonic anhydrase; CA domain;
 XX abnormal expression; neoplastic disease; cancer; gene therapy.

OS Homo sapiens.

PN W0200024913-A2.

PD 04-MAY-2000.

PF 22-OCT-1999; 99WO-US024879.

PR 23-OCT-1998; 98US-00177776.

PR 23-OCT-1998; 98US-00178115.

PR 23-OCT-1998; 98US-00178115.

PR 23-OCT-1998; 98US-00178115.

PR 23-OCT-1998; 98US-00178115.

PR 23-OCT-1998; 98US-00178115.

PR 23-OCT-1998; 98US-00178115.

(FARB) BAYER CORP.
 (VIRO-) INST VIROLOGY.
 Zavada J, Pastorekova S, Pastorek J;

The present invention describes a method of screening for
 CC preneoplastic/neoplastic disease. The method comprises: (1) determining
 CC whether abnormal MN gene expression is present in a vertebrate; and (2)

XX WP1; 2000-350752/30.
 XX A molecule which specifically binds to a site on MN protein (oncoprotein)
 XX and prevents adhesion of vertebrate cells to the protein, useful for
 XX treating preneoplastic or neoplastic diseases such as cancer.
 XX
 XX Claim 14; Page 125-126; 154pp; English.
 XX
 XX The invention relates to the inhibition of cell adhesion mediated by the
 XX MN oncoprotein (also known as the MN/CA IX isoenzyme or the MN/G250
 XX protein). The MN protein is a tumour-associated adhesion molecule which
 XX comprises a proteoglycan-like (PG) domain (AAB03017) which contains the
 XX protein's binding site, and a carbonic anhydrase (CA) domain (AAB03018).
 XX Abnormal expression of the MN protein is associated with tumorigenicity.
 XX The invention encompasses molecules (e.g., proteins and peptides) which
 XX which specifically bind to a site on the MN protein, thereby preventing
 XX adhesion of vertebrate cells to the protein in a cell adhesion assay. It
 XX also encompasses MN proteins or MN protein fragments which can be added
 XX to the extracellular environment to prevent the adhesion of vertebrate
 XX cells to each other. The invention also relates to the identification of
 XX the binding site of the MN protein and to a method of identifying a site
 XX on an MN protein to which cells adhere, comprising testing a series of
 XX overlapping peptides from the protein in a cell adhesion assay. The
 XX invention encompasses a vector comprising an expression control sequence
 XX operatively linked to a nucleic acid encoding the variable domains of a
 XX MN-specific antibody, where the domains are separated by a flexible
 XX linker peptide (AAB03035) and the vector inhibits the growth of a
 XX vertebrate preneoplastic or neoplastic cell that abnormally expresses MN
 XX protein. The invention also encompasses a vector comprising a nucleic
 XX acid encoding a cytotoxic protein or peptide operatively linked to the MN
 XX gene promoter, which inhibits the growth of a vertebrate preneoplastic or
 XX neoplastic cell. Also claimed is a repressor complex that binds to the MN
 XX gene promoter (AAB52473). MN proteins and peptides, MN-binding proteins
 XX and peptides, and expression vectors encoding such proteins and peptides
 XX are useful for treating patients with preneoplastic or neoplastic diseases
 XX (e.g., cancers) associated with or characterized by abnormal MN
 XX expression. Sequences AAB03006, AAB03015-B03022, AAB03029-B03031,
 XX AAB03036 and AAB03059-B03060 represent specific domains or regions of the
 XX human MN protein (AAB03005) referred to in the invention

Sequence 257 AA;
 SQ
 Query Match 56.5%; Score 1370; DB 3; Length 257;
 Best Local Similarity 100.0%; Pred. No. 7.8e-100; Indels 0; Gaps 0;
 Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 GDDQSHMRGDDPMPRVSPACAGRFQSPVDIRPOLAFAFCALRPLELIGFQLPPLPELR 194
 Db 1 GDDQSHMRGDDPMPRVSPACAGRFQSPVDIRPOLAFAFCALRPLELIGFQLPPLPELR 60
 QY 195 LRNGHNSVOLTPPGLEVALGGRGYRALQLHLHWGAAGRGSHYVGBRFPPEIHVHL 254
 Db 61 LRNGHNSVOLTPPGLEVALGGRGYRALQLHLHWGAAGRGSHYVGBRFPPEIHVHL 120
 QY 255 LSTAFARVDEALGRGGLAVLAALFLEBGPENSAYEQULSLREIABEGSETQVPGDIS 314
 Db 121 LSTAFARVDEALGRGGLAVLAALFLEBGPENSAYEQULSLREIABEGSETQVPGDIS 180
 QY 315 ALPSPDFSRFYFQEGSLTTPCAQGVITVFNOQYVLSAKQLHTLSLTLMGPGDSRIQLN 374
 Db 181 ALPSPDFSRFYFQEGSLTTPCAQGVITVFNOQYVLSAKQLHTLSLTLMGPGDSRIQLN 240
 QY 375 FRATQPLNGRVTEASFP 391
 Db 241 FRATQPLNGRVTEASFP 257

RESULT 15
 AAR97235
 ID AAR97235 standard; protein; 256 AA.
 XX AAR97235;
 AC AAR97235;

XX 25-MAR-2003 (revised)
 DT 26-JUL-1996 (first entry)
 XX
 DE Mtu putative oncogene MN protein carbonic anhydrase domain.
 XX
 XX Mtu; endogenous; cellular component; MN; Hela cell; diagnosis;
 KW lymphocytic choriomeningitis virus; LCMV; putative oncogene; treatment;
 KW neoplastic; pre-neoplastic; disease; antisense therapy; antibody;
 KW vaccine; vertebrate; immunisation; carbonic anhydrase.
 XX
 OS Homo sapiens.
 XX
 PN W09534650-A2.
 XX
 PD 21-DEC-1995.
 XX
 XX 15-JUN-1995; 95WO-US007628.
 PF
 XX
 PR 15-JUN-1994; 94US-00260190.
 PR 07-JUN-1995; 95US-00477504.
 PR 07-JUN-1995; 95US-00481658.
 PR 07-JUN-1995; 95US-00485049.
 PR 07-JUN-1995; 95US-00485862.
 PR 07-JUN-1995; 95US-00485863.
 PR 07-JUN-1995; 95US-00486756.
 PR 07-JUN-1995; 95US-00487077.
 XX
 PA (CIBA) CIBA CORNING DIAGNOSTICS CORP.
 PA (VIRO-) INST VIROLOGY.
 XX
 XX Zavada J, Pastorekova S, Pastorek J;
 PI
 DR WP1; 1996-049679/05.
 XX
 PT MN gene, protein and nucleic acid fragments - used as primers and probes
 PT in the detection of MN antigens and antibodies, and in the treatment of
 PT (pre)neoplastic disease.
 XX
 XX Claim 12; -; 102pp; English.
 XX
 XX The present sequence is the carbonic anhydrase domain from the protein
 XX encoded by the full length Mtu endogenous cellular component, MN, cDNA
 XX clone, which was isolated from lymphocytic choriomeningitis virus (LCMV)
 XX infected Hela cells. Persistent LCMV, the exogenous Mtu transmissible
 XX agent (MX), infection increases the expression level of the MN gene. MN
 XX is a putative oncogene, and can therefore be used in the development of
 XX prods. for the diagnosis and treatment of neoplastic (NP), or pre-NP
 XX diseases. NP diseases can be treated using DNA antisense to MN
 XX transcribed mRNA, anti-MN protein antibodies can be used for the
 XX diagnosis NP or pre-NP diseases and a vaccine contg. immunogenic amounts
 XX of the MN protein can be used to immunise a vertebrate against a NP
 XX disease associated with MN antigen expression. (Updated on 25-MAR-2003 to
 XX correct PR field.)
 XX
 SQ Sequence 256 AA;
 Query Match 56.3%; Score 1364; DB 2; Length 256;
 Best Local Similarity 100.0%; Pred. No. 2.3e-99; Indels 0; Gaps 0;
 Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 DDQSHMRGDDPMPRVSPACAGRFQSPVDIRPOLAFAFCALRPLELIGFQLPPLPELR 195
 Db 1 DDQSHMRGDDPMPRVSPACAGRFQSPVDIRPOLAFAFCALRPLELIGFQLPPLPELR 60
 QY 196 LRNGHNSVOLTPPGLEVALGGRGYRALQLHLHWGAAGRGSHYVGBRFPPEIHVHL 255
 Db 61 LRNGHNSVOLTPPGLEVALGGRGYRALQLHLHWGAAGRGSHYVGBRFPPEIHVHL 120
 QY 256 STAFARVDEALGRGGLAVLAALFLEBGPENSAYEQULSLREIABEGSETQVPGDISA 315
 Db 121 STAFARVDEALGRGGLAVLAALFLEBGPENSAYEQULSLREIABEGSETQVPGDISA 180

Qy	316	LLPDSRFFQYESSITTPCAQGVITWENQVTWLSAKQJHTLSDTLNGPDSRIQLNF	375
Db	181	LLPDSRFFQYESSITTPCAQGVITWENQVTWLSAKQJHTLSDTLNGPDSRIQLNF	240
Qy	376	RATOPPLNGRVIASFP	391
Db	241	RATOPPLNGRVIASFP	256

Search completed: March 5, 2004, 09:04:10
Job time : 61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2004, 09:02:10 ; Search time 20 Seconds

(without alignments)
2207.596 Million cell updates/sec

Title: US-09-967-237-2

Sequence: 1 MAPLCPSPMLPILIPAPAPG.....RRGTGVSYPAPAEVETGA 459

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match Length	DB ID	Description
1	2424	100.0	459 2	I38013 p54/58N - human
2	526	21.7	307 2	A23993 carbonate dehydrat
3	511	21.1	319 2	S71877 carbonate dehydrat
4	492	20.3	308 1	CRH06 carbonate dehydrat
5	425	17.5	303 2	S12867 carbonate dehydrat
6	412	17.0	290 2	UN0576 carbonate dehydrat
7	409	16.9	264 1	CRH07 carbonate dehydrat
8	402.5	16.6	260 2	T08463 carbonate dehydrat
9	391	16.1	259 1	CRB02 carbonate dehydrat
10	380.5	15.7	260 1	CRMS2 carbonate dehydrat
11	379.5	15.7	260 1	CRH02 carbonate dehydrat
12	379	15.6	259 1	CRSH2 carbonate dehydrat
13	374.5	15.4	259 1	CRB02 carbonate dehydrat
14	368.5	15.2	260 2	UH0527 carbonate dehydrat
15	366.5	15.1	260 1	CRM01R carbonate dehydrat
16	364.5	15.0	260 2	JC2580 carbonate dehydrat
17	363	15.0	261 2	A26344 carbonate dehydrat
18	362.5	15.0	261 2	UN0836 carbonate dehydrat
19	359.5	14.8	260 2	I52551 carbonate dehydrat
20	359.5	14.8	261 1	CRH01 carbonate dehydrat
21	359.5	14.8	261 1	UN0835 carbonate dehydrat
22	354	14.6	260 1	CRH01D carbonate dehydrat
23	354	14.6	260 2	A43641 carbonate dehydrat
24	347	14.3	309 2	A51900 carbonic anhydrase
25	343.5	14.2	260 1	CRH03 carbonate dehydrat
26	342	14.1	304 2	A59261 carbonate dehydrat
27	338.5	14.0	259 2	A22612 carbonate dehydrat
28	333	13.7	334 2	T16772 hypochlorite prote
29	328.5	13.6	305 1	CRH05 carbonate dehydrat

30	326.5	13.5	1442 1	B46148 protein-lysine-p
31	325	13.4	298 2	S12579 carbonate dehydrat
32	323	13.3	1445 1	A46148 protein-lysine-p
33	322	13.3	246 2	T16575 hypochlorite prote
34	321	13.2	235 2	A22962 carbonate dehydrat
35	304.5	12.6	2314 1	A46151 protein-lysine-p
36	303	12.5	1422 2	T42636 protein-lysine-p
37	301	12.4	306 2	A46993 protein-lysine-p
38	288.5	11.9	312 1	CRH04 carbonate dehydrat
39	288.5	11.3	319 2	T15137 protein-lysine-p
40	273.5	11.3	275 2	H83694 carbonate dehydrat
41	270.5	11.2	324 2	T14196 carbonic anhydrase
42	268.5	11.1	259 2	T12849 hypochlorite prote
43	264.5	10.9	328 2	UB0375 carbonic anhydrase
44	263.5	10.9	239 2	F82479 carbonic anhydrase
45	261.5	10.8	365 2	F88449 protein F54D8.4 [1

ALIGNMENTS

RESULT 1

I38013
p54/58N - human
C/Species: Homo sapiens (man)
C/Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-Nov-1999
C/Accession: I38013
R/Pastorek, U.; Pastorekova, S.; Callebaut, I.; Mornon, J.; Zelnik, V.; Opavsky, R.; Oncogene 9, 2877-2888, 1994
A/Title: Cloning and characterization of MN, a human tumor-associated protein with a A/Reference number: I38013; MUID:94366734; PMID:8084592
A/Accession: I38013
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-459 <RES>
A/Cross-references: EMBL:X66839; NID:G1000701; FIDN:CAA47315.1; PID:G1000702
A/Genetics:
A/Gene: MATUN
C/Superfamily: carbonic anhydrase homology
F/141-390/Domain: carbonic anhydrase homology <CNA>

Query Match	100.0%;	Score 2424;	DB 2;	Length 459;
Best Local Similarity	100.0%;	Pred. No. 4.1e-147;		
Matches 459;	Conservative	0;	Mismatches 0;	Indels 0;
0;	Gaps 0;			
QY	1	MAPLCPSPMLPILIPAPAPGLTVQLLSILLMPVHPQRLPRNOEDSPLGGSGGEDDPL	60	
DB	1	MAPLCPSPMLPILIPAPAPGLTVQLLSILLMPVHPQRLPRNOEDSPLGGSGGEDDPL	60	
QY	61	GSEDLPSEEDSFRSEDPGSEEDLPGEEDLPGEEDLPVKKPKSEEGSLKLEDLPTVAPG	120	
DB	61	GSEDLPSEEDSFRSEDPGSEEDLPGEEDLPGEEDLPVKKPKSEEGSLKLEDLPTVAPG	120	
QY	121	DQEPONNAHRDKEDDQSHRWYGGDPMPWPVSACAGRPSPVDIRPOLAFCPALRPL	180	
DB	121	DQEPONNAHRDKEDDQSHRWYGGDPMPWPVSACAGRPSPVDIRPOLAFCPALRPL	180	
QY	181	ELLGQLPPLPELRIRNGHSGVQLTLPGLMALGPGEYRRAQLHLHWAAGPQSEHT	240	
DB	181	ELLGQLPPLPELRIRNGHSGVQLTLPGLMALGPGEYRRAQLHLHWAAGPQSEHT	240	
QY	241	VEGHFPFAIHVHLSFAFVDEALGRPGSLATLAFLBEGPEENSAYEQLLSRLEEA	300	
DB	241	VEGHFPFAIHVHLSFAFVDEALGRPGSLATLAFLBEGPEENSAYEQLLSRLEEA	300	
QY	301	EEGSETQVPGDISALPSDSRYFQYEGSLTPPCAGVITVFQNTWLSAKQMTLS	360	
DB	301	EEGSETQVPGDISALPSDSRYFQYEGSLTPPCAGVITVFQNTWLSAKQMTLS	360	
QY	361	DTLWPGSPRLQINPRALQPLNGVITASFPAGVDSPPRAAPQVNSCAAGDITLAVF	420	
DB	361	DTLWPGSPRLQINPRALQPLNGVITASFPAGVDSPPRAAPQVNSCAAGDITLAVF	420	

QY 421 GILFATSVAFIVQNRQRGRGTGGVSVRPAEVAETGA 459
 DB 421 GILFATSVAFIVQNRQRGRGTGGVSVRPAEVAETGA 459

RESULT 2

A29993
 C:Enzyme: carbonic dehydratase (EC 4.2.1.1) VI - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C>Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 23-May-1997
 C/Keywords: A29993
 R/Fernley, R.T.; Wright, R.D.; Coghlan, J.P.
 Biochemistry 27, 2815-2820, 1988
 A>Title: Complete amino acid sequence of ovine salivary carbonic anhydrase.
 A/Reference number: A29993; PMID:88294021; PMID:3135834
 A/Accession: A29993
 A/Molecule type: protein
 A/Residues: 1-307 <FER>

A/Experimental source: parotid gland
 A/Note: 63-Mer and 297-11e were also found
 C/Superfamily: carbonic dehydratase; carbonic anhydrase homology
 C/Keywords: carbon-oxygen lyase; hydro-lyase; zinc
 F:6-261/Domain: carbonic anhydrase homology <CAH>

Query Match 21.7%; Score 526; DB 2; Length 307;
 Best Local Similarity 43.2%; Pred. No. 1.9e-26;
 Matches 111; Conservative 40; Mismatches 98; Indels 8; Gaps 4;

QY 141 WRG---GDPMPRSPACAGRPSPVDIRPOLAFCPALRPLELGGFQLPPLRLR 196
 DB 6 WYSEGMDEAMPLEPKCGGRSPDLOKMKYQVSRNALNTGYGLM-GEFPVT 64
 QY 197 NNGHSVQLTPPGLEMLGPGREYRALQLHLHWGAAGR--PGSEHTVEGHRPAEIHVH 254
 DB 65 NNGHTVQISLPSTMTSTSDGTQYLAKQMFHWGASSEISGSEHTVDGKRYVEIHVH 124
 QY 255 LSTAFARVDEALGPRGGLAVLAFL-EGPEENSAYEQLLSRLSEIAEGSEFQVPLDI 313
 DB 125 YNSKNSYEAKQEPDGLVALVEKDYENAYYSFISHLDIRYAGQSTVLRGLDI 184
 QY 314 SALLPSDFSRFYQEGSLTTPPCAGQIVTWFNQTMLSAKQLHTLSDTLMGPGDSRLQL 373
 DB 185 EDLPGDRLRYYSYSLTTPPCSTENVMFVADTVKLSKQVEKLENSLHNQNTION 244
 QY 374 NFRATQPLNGRVIEASF 390
 DB 245 DYRRTQPLNHRVVEANF 261

RESULT 3

S71877
 C:Enzyme: carbonic dehydratase (EC 4.2.1.1) isoform VI precursor - bovine
 N/Alternate names: carbonic anhydrase VI
 C/Species: Bos primigenius taurus (cattle)
 C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 20-Jun-2000
 C/Accession: S71877; S03862
 R/Jiang, W.; Mollach, J.T.; Gupta, D.
 Biochem. J. 318, 291-296, 1996
 A>Title: Sequence of bovine carbonic anhydrase VI: potential recognition sites for N-ac
 A/Reference number: S71877; PMID:96558528; PMID:8761494
 A/Accession: S71877
 A/Molecule type: mRNA
 A/Residues: 1-319 <UIA>

A/Cross-references: EMBL:X96503; NID:g1526571; PIDN:CAA65357.1; PID:g1526572
 A/Experimental source: submandibular gland
 R/Fernley, R.T.; Darling, P.; Aldred, P.; Wright, R.D.; Coghlan, J.P.
 Biochem. J. 259, 91-96, 1989
 A>Title: Tissue and species distribution of the secreted carbonic anhydrase isoenzyme.
 A/Reference number: S03862; PMID:89246311; PMID:2497732
 A/Accession: S03862
 A/Molecule type: protein
 A/Residues: 15,'S',17-36,'X',38-39 <FER>
 C/Genetics:

A/Gene: CAH6
 C/Function:
 A/Description: catalyzes the reversible dissociation of carbonic acid to carbon diox
 C/Superfamily: carbonic dehydratase; carbonic anhydrase homology
 C/Keywords: carbon-oxygen lyase; glycoprotein; hydro-lyase; metalloprotein; zinc
 F:1-14/Domain: signal sequence #status predicted <SIG>
 F:15-319/Product: carbonic dehydratase isoform VI #status experimental <MAT>
 F:18-273/Domain: carbonic anhydrase homology <CAH>
 F:37-219/Disulfide bonds: #status predicted
 F:62,251/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:106,108,133/Binding site: zinc (His) #status predicted

Query Match 21.1%; Score 511; DB 2; Length 319;
 Best Local Similarity 42.0%; Pred. No. 1.8e-25;
 Matches 108; Conservative 40; Mismatches 97; Indels 12; Gaps 4;

QY 137 DQSHMRVGGDPMPRSPACAGRPSPVDIRPOLAFCPALRPLELGGFQLPPLRLR 196
 DB 26 DEKHRL-----QYDCCGTRQSPIDLMKKVRYPPLRALNTGYGLRQ-GEFPMT 76
 QY 197 NNGHSVQLTPPGLEMLGPGREYRALQLHLHWGAAGR--PGSEHTVEGHRPAEIHVH 254
 DB 77 NNGHTVQISLPSSNRMSTSDGQTLAKQMFHWGASSEISGSEHTVDGKRYVEIHVH 136
 QY 255 LSTAFARVDEALGPRGGLAVLAFL-EGPEENSAYEQLLSRLSEIAEGSEFQVPLDI 313
 DB 137 YHSKYSYEAKQEPDGLVALVEKDYENAYYSNFIHLDIRYAGQSTVLRGLDI 196
 QY 314 SALLPSDFSRFYQEGSLTTPPCAGQIVTWFNQTMLSAKQLHTLSDTLMGPGDSRLQL 373
 DB 197 QDMPLGDRYYSYSLTTPPCSTENVMFVADTVKLSKQVEKLENSLHNQNTION 256
 QY 374 NFRATQPLNGRVIEASF 390
 DB 257 NYRSTQPLNHRVVEANF 273

RESULT 4

CRT06
 C:Enzyme: carbonic dehydratase (EC 4.2.1.1) VI precursor - human
 N/Alternate names: carbonic anhydrase VI; salivary carbonic anhydrase; secreted carb
 C/Species: Homo sapiens (man)
 C>Date: 31-May-1991 #sequence_revision 05-May-1995 #text_change 02-Sep-1997
 C/Accession: A37917
 R/Aldred, P.; Fu, P.; Barrett, G.; Penschow, J.D.; Wright, R.D.; Coghlan, J.P.; Fern
 Biochemistry 30, 569-575, 1991
 A>Title: Human secreted carbonic anhydrase: cDNA cloning, nucleotide sequence, and h
 A/Reference number: A37917; PMID:9105141; PMID:1899030
 A/Accession: A37917
 A/Molecule type: mRNA
 A/Residues: 1-308 <AUD>

A/Cross-references: GB:M57892; GB:J05305
 A/Experimental source: salivary gland
 A/Note: the authors translated the codon GAG for residue 248 as Gln
 C/Genetics:

A/Gene: GDB:CA6
 A/Cross-references: GDB:125350; OMIM:114780
 A/Map position: 1936-1936
 C/Function:

A/Description: catalyzes the reversible dissociation of carbonic acid to carbon diox
 C/Superfamily: carbonic dehydratase; carbonic anhydrase homology
 C/Keywords: carbon-oxygen lyase; glycoprotein; hydro-lyase; metalloprotein; pyroglyut
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-108/Product: carbonic dehydratase VI #status predicted <MAT>
 F:23-278/Domain: carbonic anhydrase homology <CAH>
 F:118/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predi
 F:442-224/Disulfide bonds: #status predicted
 F:67,256/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:111,115,138/Binding site: zinc (His) #status predicted

Query Match 20.3%; Score 492; DB 1; Length 308;
 Best Local Similarity 39.6%; Pred. No. 2.8e-24;

Matches 103; Conservative 44; Mismatches 105; Indels 6; Gaps 4;

QY 139 SHRRYQ---GDPWRVSDACAGRFQSPVDIRPOLAFCALPRLLELIGQLPELRL 194

Db 21 SDMTYBEGALDEAHMQRHPACGCGQSPINLQRTKVRNPNLSKLNMTGVE--CQAGSFP 79

QY 195 LRNNGHSVQTLTPPGEMALGPGRERFALQHLHWAGAGR--PGEHTEVGHRRPAEIHV 252

Db 80 MANNHGHVQIGLPTSMRTMYADQIVYIAQGMFHHGASSETISGSEHTVQDGRHYEIH 139

QY 253 VHSFAFVDAALGPPGGLAVIAFLF-EGEENSAYEQLLSRLSEIAEGSETOVPG 311

Db 140 VHNYSKRYKTDIQDAPDGLAVLAFAVEYKVPENTYVSNFISHLANIKYPGQRTLTGL 199

QY 312 DISALLPSCDSRYFOYEGSLTPPCAGQIVMTVFNQTLWLSAKOHLTSLDLMRGSR 371

Db 200 DVQDMLPRNQHHTYTHGSLTTPCTENVMHFLADFYKLSRTQVWKLNSLDHRNKT 259

QY 372 QLNFRATOPUNGIVIASFP 391

Db 260 HNDYRRTQPLNHRVSNFP 279

RESULT 5

S12867
carbonate dehydratase-related protein - mouse
N/Alternate names: carbonic anhydrase-related protein
C/Species: Mus musculus (house mouse)
C/Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 22-Jun-1999
C/Accession: S12867
R/Kato, K.
FEBS Lett. 271, 137-140, 1990
A/Title: Sequence of a novel carbonic anhydrase-related polypeptide and its exclusive pH
A/Reference number: S12867; MUID:91032050; PMID:121526
A/Accession: S12867
A/Molecule type: mRNA
A/Residues: 1-303 <KAT>
A/Cross-references: EMBL:X61397; NID:950286; PIDN:CAA3668.1; PID:950287
C/Superfamily: carbonate dehydratase; carbonic anhydrase homology
F/30-290/Domain: carbonic anhydrase homology <CAH>

Query Match 17.5%; Score 425; DB 2; Length 303;
Best Local Similarity 32.9%; Pred. No. 5e-20;
Matches 104; Conservative 61; Mismatches 119; Indels 32; Gaps 10;

QY 110 LEDLPTVE-APGDPQPNNAHRDQSDQSHWRYGDPMPRVSPACAGRFQSPVDIRP 168

Db 1 MADLSFIEDVAFAPEKEDEDEEEDEEVE--WGVEEGVEMGLVFPDANGRYQSPINLNS 57

QY 169 QLAFCPALRPLELIGFOLP---LPELRLRNNGHSVQTLTPPGLEMALGP---GREYR 221

Db 58 REARVDP-----LDVRLSPYVVCRCCEVTNDGHTIQVILKSKSVLSGGLPQGGFE 112

QY 222 ALQHLHWAGAGRPGSEHTVEGHRPAEIHVHL-STAFARVDAALGPPGLAVLAFLF 280

Db 113 LVEVFFHWGRNQSRSEHTVFKAPFMELHLIHNMTSLFGSIDAVKPGHAIALFVQ 172

QY 281 EGPEENSAVEQLLSRLSEIAEGSETOVPGDISALLPSDSR-YFQYEGSLTPPCAG 339

Db 173 IG-KEHVGLKAVTELLQDIQYKSKTIPCNPNVILLPDLRLRYWYEGSLTTPCSBG 231

QY 340 VMTVFNQTLWLSAKOL-----HTLSDTLWPGDSRLQLNFRATOPUNGIVIASFP 391

Db 232 VMTLFRYPLTISQIIEFRRLRTHVKGALVEGCGIIGDNRFPQPLSDRVIRAAF 290

QY 392 AGVDSPPRAEYVQIN 407

Db 291 ----SSQDRREQTHL 302

RESULT 6
JN0576
carbonate dehydratase (EC 4.2.1.1) VIII - human

N/Alternate names: carbonic anhydrase-related protein (CARP)
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
C/Accession: JN0576
R/Skaggs, L.A.; Bergenhem, N.C.H.; Venta, P.J.; Taahian, R.E.
Gene 126, 291-292, 1993
A/Title: The deduced amino acid sequence of human carbonic anhydrase-related protein
A/Reference number: JN0576; MUID:93246262; PMID:8482548
A/Accession: JN0576
A/Molecule type: mRNA
A/Residues: 1-290 <SKA>
A/Cross-references: GB:L04656; NID:9179937; PIDN:AAA3653.2; PID:95069431
C/Genetic:
A/Genes: GDB:CA8; CALS; CARP
A/Cross-references: GDB:141839; OMIM:114815
A/Map position: 8pter-8qter
C/Function:
A/Description: catalyzes the reversible dissociation of carbonic acid to carbon dioxide
A/Note: this protein has not been demonstrated to be enzymatically active
C/Superfamily: carbonate dehydratase; carbonic anhydrase homology
C/Keywords: carbon-oxygen lyase; hydro-lyase; metalloprotein; zinc
F/2-261/Product: carbonate dehydratase VIII #status predicted <VAT>
F/29-289/Domain: carbonic anhydrase homology <CAH>

Query Match 17.0%; Score 412; DB 2; Length 290;
Best Local Similarity 32.9%; Pred. No. 3.2e-19;
Matches 98; Conservative 57; Mismatches 117; Indels 26; Gaps 8;

QY 110 LEDLPTVAAPGDPQPNNAHRDQSDQSHWRYGDPMPRVSPACAGRFQSPVDIRP 169

Db 1 MADLSFIEDVAFAPEKEDEDEEEDEEVE--WGVEEGVEMGLVFPDANGRYQSPINLNS 57

QY 170 LAAPCALRPLELIGFOLP---LPELRLRNNGHSVQTLTPPGLEMALGP---GREYR 222

Db 58 EARVDP-----LDVRLSPYVVCRCCEVTNDGHTIQVILKSKSVLSGGLPQGGFE 112

QY 223 LQHLHWAGAGRPGSEHTVEGHRPAEIHVHL-STAFARVDAALGPPGLAVLAFLF 281

Db 113 YEVRHWGRNQSRSEHTVFKAPFMELHLIHNMTSLFGSIDAVKPGHAIALFVQ 172

QY 282 EGPEENSAVEQLLSRLSEIAEGSETOVPGDISALLPSDSR-YFQYEGSLTPPCAG 340

Db 173 G-KEHVGLKAVTELLQDIQYKSKTIPCNPNVILLPDLRLRYWYEGSLTTPCSBG 231

QY 341 TMTVFNQTLWLSAKOL-----HTLSDTLWPGDSRLQLNFRATOPUNGIVIASF 390

Db 232 TMTLFRYPLTISQIIEFRRLRTHVKGALVEGCGIIGDNRFPQPLSDRVIRAAF 289

RESULT 7

CRH07
carbonate dehydratase (EC 4.2.1.1) VII - human
N/Alternate names: carbonic anhydrase VII
C/Species: Homo sapiens (man)
C/Date: 10-Feb-1995 #sequence_revision 05-May-1995 #text_change 18-Jun-1999
C/Accession: A55272
R/Montgomery, J.C.; Venta, P.J.; Eddy, R.L.; Fukushima, Y.S.; Showne, T.B.; Taahian, R.
Genomics 11, 835-848, 1991
A/Title: Characterization of the human gene for a newly discovered carbonic anhydrase
A/Reference number: A55272; MUID:92147127; PMID:1783592
A/Accession: A55272
A/Molecule type: DNA
A/Residues: 1-264 <MON>
A/Cross-references: GB:MT6423; NID:9179964; PIDN:AAA5193.1; PID:9179967
A/Note: sequence extracted from NCBI backbone (NCBIN:80199, NCBIN:80201, NCBIN:80205,
C/Genetic:
A/Genes: GDB:CA7
A/Cross-references: GDB:119741; OMIM:114770
A/Map position: 16q22.1-16q22.1
A/Introns: 14/1; 80/1; 119/3; 151/3; 172/3; 224/3
C/Function:
A/Description: catalyzes the reversible dissociation of carbonic acid to carbon dioxide
A/Note: this form is expressed in salivary gland and other tissues

C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
 C:Keywords: carbon-oxygen lyase; hydro-lyase; metalloprotein; zinc
 F/4-264/Product: carbonate dehydratase VII #status predicted <Mat>
 F/7-262/Domain: carbonic anhydrase homology <CAH>
 F/56/58,121/Binding site: zinc (His) #status predicted

Query Match 16.9%; Score 409; DB 1; Length 264;
 Best Local Similarity 37.5%; Pred. No. 1,4e-19;
 Matches 101; Conservative 41; Mismatches 99; Indels 28; Gaps 11;

QY 141 WRYG---GDPWPRVSPACAGRFQSPVDIRPOLAFCPALPPELLGFOPLPPELRN 197
 DB 7 WYGGDDGSHMKLYPIAQGRQSPINISSQAVYSPSLQPLEL---SYEACMSLSTIN 63
 QY 198 NCHSVOL-----TLPRG--LEWALGGRREYALQHLHMGAAAGPSEHNYEGRFP 247
 DB 64 NCHSVQVDPNDSDDTVTGPLE---GP---YIKQPHFMGKKHVDVGEHVDGKSPF 117
 QY 248 AEIHYVHLST-AFAVDEALGRPGGLAVLAFLLEGPEENSAVQLRLLEIAEGSEET 306
 DB 118 SELHYVMWAKKYSTFGAASAPDGLAVGVFLETG-DEHSMNRLTDALYVVRPKGTXA 176
 QY 307 QYRGDISALPSDSRYFOYEGSLTTPPCAGQVITWYFNQVTLMSKQHLTSLTM-G 365
 DB 177 QSCFNPCKLPLAS-RHYWTYPSLTTPLESEVTVIWRPILCISERQNGKFRSLFTS 235
 QY 366 PGDSRLQL--NFRATQPLNGRVLEASFP 392
 DB 236 EDDERIMWNNRPPQPLKGRVYKASFP 264

RESULT 8

T08463
 Carbonate dehydratase (EC 4.2.1.1) - zebra fish
 C:Species: Brachydanio rerio (zebra fish)
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
 C:Accession: T08463
 R:Peterson, R.E.; Tu, C.; Linser, P.J.
 U:MOJ. Evol. 44, 432-439, 1997
 A>Title: Isolation and characterization of a carbonic anhydrase homologue from the zebra
 A:Reference number: Z16422; MUID:97250269; PMID:9089083
 A:Accession: T08463
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-260 <PRT>
 A:Cross-references: EMBL:U55177; NID:92576334; PID:92576335
 C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
 C:Keywords: carbon-oxygen lyase; hydro-lyase; zinc
 F/5-259/Domain: carbonic anhydrase homology <CAH>

Query Match 16.6%; Score 402.5; DB 2; Length 260;
 Best Local Similarity 35.8%; Pred. No. 1.1e-18;
 Matches 95; Conservative 37; Mismatches 108; Indels 25; Gaps 7;

QY 141 WRYG---GDPWPRVSPACAGRFQSPVDIRPOLAFCPALPPELLGFOPLPPELRN 197
 DB 5 WYGGADGPFESMAESFFPIANGPROSPIDIVTQHPDSIKHKL---KIDPATKSLIN 61
 QY 198 NCHSVOL-----TLPRG--LEWALGGRREYALQHLHMGAAAGPSEHNYEGRFP 248
 DB 62 NCHSVQVDPNDSDDTVTGPLE---GP---YIKQPHFMGKKHVDVGEHVDGKSPF 116
 QY 249 EIHVHLSTAFARVDEALGRPGGLAVLAFLLEGPEENSAVQLRLLEIAEGSEETV 308
 DB 117 EHLHYVMWAKKYSTFGAASAPDGLAVGVFLETG-AANPLRLQVLDALDIKSKGQTF 175
 QY 309 PGDISALPSDSRYFOYEGSLTTPPCAGQVITWYFNQVTLMSKQHLTSLTM---G 365
 DB 176 ANPDKTLPLASD-YWYBGSLLTPPLLSVTVIKETISVSPQMAKFRSLFSSSG 234
 QY 366 PGDSRLQLNFRATQPLNGRVLEASFP 390
 DB 235 ETPCCMVNRRPPQPLKGRKVRASF 259

RESULT 9
 CRB02
 Carbonate dehydratase (EC 4.2.1.1) II - bovine (tentative sequence)

N:Alternate names: carbonic anhydrase II
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 31-Mar-2000
 C:Accession: A01144
 R:Schacky, M.; Limozin, N.; Filippi-Foveau, D.; Guilian, J.M.; Laurent-Tabusse, G.
 Biochimie 58, 1071-1082, 1976
 A>Title: Structure primaire de l'anhydrase carbonique erythrocytaire bovine CI. II.
 A:Reference number: A90669; MUID:77065798; PMID:826282
 A:Accession: A01144

A:Molecule type: protein
 A:Residues: 1-259 <SCI>
 R:Guilian, J.M.; Limozin, N.; Mallet, B.; Di Costanzo, J.; Charrel, M.
 Biochimie 59, 293-302, 1977
 A>Title: Independance genetique de deux formes de l'anhydrase carbonique erythrocytaire
 A:Reference number: A90672; MUID:77242599; PMID:19093
 A:Contents: annotation

A>Note: One minor and two major forms were isolated chromatographically. One of the
 C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
 C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; zinc
 F/4-258/Domain: carbonic anhydrase homology <CAH>
 F/1/Modified site: acetylated amino end (Ser) #status experimental
 F/93,95,118/Binding site: zinc (His) #status predicted

Query Match 16.1%; Score 391; DB 1; Length 259;
 Best Local Similarity 35.5%; Pred. No. 6.1e-18;
 Matches 93; Conservative 43; Mismatches 110; Indels 16; Gaps 8;

QY 140 HMRYG---GDPWPRVSPACAGRFQSPVDIRPOLAFCPALPPELLGFOPLPPELRN 196
 DB 3 HMGVKGKBPZHMHKKPFPIANGRGSPVINDTAVVQDPLALPLV---YGEATSRNV 59
 QY 197 NCHSVOLTLPRG--LEWALGGRREYALQHLHMGAAAGPSEHNYEGRFP 247
 DB 60 NCHSVQVDPNDSDDTVTGPLE---GP---YIKQPHFMGKKHVDVGEHVDGKSPF 119
 QY 253 VHLSTAFARVDEALGRPGGLAVLAFLLEGPEENSAVQLRLLEIAEGSEETVPGD 312
 DB 120 VHMWTKGDPGTAAQPDGLAVGVFLKYG-DANPLRLQVLDALDIKSKGSTOFPNPD 178
 QY 313 ISALPSDSRYFOYEGSLTTPPCAGQVITWYFNQVTLMSKQHLTSLTM---G 369
 DB 179 PGSLPLPVLD-YWYBGSLLTPPLLSVTVIKETISVSPQMAKFRSLFSSSG 237
 QY 370 RLQINFRATQPLNGRVLEASFP 391
 DB 238 LMLANRRPPQPLKGRKVR-GFP 258

RESULT 10

CRMS2
 Carbonate dehydratase (EC 4.2.1.1) II - mouse
 N:Alternate names: carbonic anhydrase II
 C:Species: Mus musculus (house mouse)
 C>Date: 25-Feb-1985 #sequence_revision 06-Feb-1995 #text_change 18-Jun-1995
 C:Accession: A23900; B23202; A01143; A20535; I51949
 R:Ventra, P.J.; Montgomerie, J.C.; Hewett-Emmett, D.; Wiebauer, K.; Tashian, R.E.
 J. Biol. Chem. 260, 12130-12135, 1985
 A>Title: Structure and exon to protein domain relationships of the mouse carbonic an
 A:Reference number: A23900; MUID:86008276; PMID:2995362
 A:Accession: A23900
 A:Molecule type: DNA
 A:Residues: 1-260 <VEN>
 A:Experimental source: strain YBR
 R:Ventra, P.J.; Montgomerie, C.; Hewett-Emmett, D.; Tashian, R.E.
 Biochim. Biophys. Acta 826, 195-201, 1985
 A>Title: Comparison of the 5' regions of human and mouse carbonic anhydrase II genes
 A:Reference number: A90655; MUID:86077780; PMID:3000449
 A:Accession: B23202

A/Molecule type: DNA
 A/Residues: 1-77 <VE2>
 A/Note: the authors translated the codon CAG for residue 39 as His
 R/Curtis, P.J.; Wilthers, E.; Demuth, D.; Wett, R.; Venter, P.J.; Tashian, R.E.
 Gene 25, 325-332, 1993
 A/Title: The nucleotide sequence and derived amino acid sequence of cDNA coding for mouse
 A/Reference number: A01143; MUID:84109569; PMID:6420240
 A/Accession: A01143
 A/Molecule type: mRNA
 A/Residues: 2-38, 'H', '40-260 <CUR>
 A/Cross-references: GB:K00812; GB:M1830; NID:G192333; PIDN:AAA7336.1; PID:G
 A/Note: initiator Met not shown
 R/Curtis, P.J.
 J. Biol. Chem. 258, 4459-4463, 1983
 A/Title: Cloning of mouse carbonic anhydrase mRNA and its induction in mouse erythrocyte
 A/Reference number: A20539; MUID:83161023; PMID:6187736
 A/Accession: A20539
 A/Molecule type: mRNA
 A/Residues: 155-178, 214-240 <CU2>
 R/Venter, P.J.; Montgomery, J.C.; Wiebauer, K.; Hewett-Emmett, D.; Tashian, R.E.
 Ann. N. Y. Acad. Sci. 429, 309-323, 1984
 A/Title: Organization of the mouse and human carbonic anhydrase II genes.
 A/Reference number: I51949; MUID:84255152; PMID:6331255
 A/Accession: I51949
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 241-260 <RES>
 A/Cross-references: GB:M25944; NID:G199078; PIDN:AAA39505.1; PID:G199079
 A/Accession: M25944
 A/Status: Car-2
 C/Genetics:
 A/Introns: 12/1, 78/1, 117/3, 144/1, 169/3, 221/3
 A/Superfamily: carbonate dehydratase; carbonic anhydrase homology
 C/Keywords: carbon-oxygen lyase; hydro-lyase; zinc
 F/5-285/Domain: carbonic anhydrase homology <CAH>
 F/4, 96, 119/Binding site: zinc (His) #status predicted

Query Match 15.7%; Score 380.5; DB 1; Length 260;
 Best Local Similarity 35.6%; Pred. No. 2,8e-17;
 Matches 93; Conservative 41; Mismatches 112; Indels 15; Gaps 6;

QY 140 HRRY--GGDPWPRVSPACAGRFSPVDIRPQLAAFCFPAIRPLELLGFGQLPPLPELRRLR 196
 DB 4 HWGSKHNGPENWMDDFPANGDRQSPVDIDTAQHDPALQCP--LISYD--KAASKSIY 60
 QY 197 NNGHSYQVLTLPGLFEMAL--GP--GREYALQHLHMGACRPGSEHTEGHPPEAIEHY 252
 DB 61 NNGHSFNVBFDDSCQNAVILKGGPLSDSYRLQFHFHMOSSDQGSSEHTVKKKYALEHL 120
 QY 253 VILSTAFARVDEALRPGGLAVLAALFEEGPEENSAYEQILSRLEIAEGSETOVPGLD 312
 DB 121 VMNNTKYGDFGKAVQDPGLAVLGYFLKIGP--ASQGGKVEALHSIKTKGKAAPANFD 179
 QY 313 ISALLPSDFSRFYQYEGSLTTPPCAGYIWTVPQOTWLSAKOL--HTLSDTLMPGDS 369
 DB 180 PGLLPGNLD--YMTYPSGLTTPPLLECCTWLVLEEPITVSSEQMSHRTFLTNVEEDGAE 238
 QY 370 RIQLNFRATQPLNGEIVASFP 390
 DB 238 AMVDNMRPAQPLNKRKIKASF 259

RESULT 11
 CRU2
 carbonate dehydratase (EC 4.2.1.1) II [validated] - human
 N/A/Alternate names: carbonic anhydrase II; hepatic carbonic anhydrase
 C/Species: Homo sapiens (man)
 C/Date: 07-May-1981 #sequence revision 05-May-1995 #text change 15-Sep-2000
 C/Accession: A27175; A23202; A92147; I37214; I51663; I51877; A01141
 R/Murakami, H.; Marelich, G.P.; Grubb, J.H.; Kyle, J.W.; Sly, W.S.
 Genomics 1, 159-166, 1987
 A/Title: Cloning, expression, and sequence homologies of cDNA for human carbonic anhydrase
 A/Reference number: A27175; MUID:86085190; PMID:3121456
 A/Accession: A27175

A/Molecule type: mRNA
 A/Residues: 1-260 <MR>
 A/Cross-references: GB:J03037; NID:G179771; PIDN:AAA51908.1; PID:G179772
 R/Venter, P.J.; Montgomery, C.; Hewett-Emmett, D.; Tashian, R.E.
 Biochim. Biophys. Acta 826, 195-201, 1985
 A/Title: Comparison of the 5' regions of human and mouse carbonic anhydrase II genes
 A/Reference number: A90655; MUID:86077780; PMID:3000449
 A/Accession: A23202
 A/Molecule type: DNA
 A/Residues: 1-77 <VEN>
 A/Cross-references: GB:X03251; GB:M18100; GB:M77181; NID:G179778; PIDN:AAA51909.1; PI
 R/Henderson, L.E.; Henriksson, D.; Nyman, P.O.
 J. Biol. Chem. 251, 5457-5463, 1976
 A/Title: The primary structure of human carbonic anhydrase C.
 A/Reference number: A92194; MUID:77060709; PMID:823150
 A/Accession: A92194
 A/Molecule type: protein
 A/Residues: 2-260 <HEN>
 R/Uhl, K.T.D.; Deutsch, H.F.
 J. Biol. Chem. 249, 2329-2337, 1974
 A/Title: Human carbonic anhydrases. XII. The complete primary structure of the C iso
 A/Reference number: A92147; MUID:74143468; PMID:4207120
 A/Accession: A92147
 A/Molecule type: protein
 A/Residues: 2-260 <LIN>
 R/Montgomery, J.C.; Venter, P.J.; Tashian, R.E.; Hewett-Emmett, D.
 Nucleic Acids Res. 15, 4687, 1987
 A/Title: Nucleotide sequence of human liver carbonic anhydrase II cDNA.
 A/Reference number: I37214; MUID:87231043; PMID:3108657
 A/Accession: I37214
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-260 <RE3>
 A/Cross-references: EMBL:Y00339; NID:G29586; PIDN:CAA68426.1; PID:G29587
 A/Experimental source: liver
 A/Note: submitted to the EMBL/Genbank/DBJ databases by David Hewett-Emmett 01-JUL-19
 Am. J. Hum. Genet. 49, 1082-1090, 1991
 A/Title: Carbonic anhydrase II deficiency syndrome in a Belgian family is caused by a
 e.
 A/Reference number: I51863; MUID:92026087; PMID:1928091
 A/Accession: I51863
 A/Status: translation not shown; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-260 <RES>
 A/Cross-references: GB:M77181; NID:G179778; PIDN:AAA51909.1; PID:G179780
 A/Note: the complete nucleotide sequence is not shown
 R/Hu, P.Y.; Ernst, A.R.; Sly, W.S.; Venter, P.J.; Skaggs, L.A.; Tashian, R.E.
 Am. J. Hum. Genet. 54, 602-608, 1994
 A/Title: Carbonic anhydrase II deficiency: single-base deletion in exon 7 is the pred
 A/Reference number: I51871; MUID:94175074; PMID:8128957
 A/Accession: I51871
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 225-242 <RE2>
 A/Cross-references: GB:G69526; NID:G545850; PIDN:AAA30170.1; PID:G545851
 A/Note: wild type shown; mutant contains frameshift after residue 226
 R/Eriksson, A.B.; Jones, T.A.; Liljas, A.
 submitted to the Brookhaven Protein Data Bank, February 1989
 A/Reference number: A50085; PDB:1CA2
 A/Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 4-259
 R/Liljas, A.; Kannan, K.K.; Bergsten, P.C.; Waara, I.; Fridborg, K.; Strandberg, B.;
 Nature New Biol. 235, 131-137, 1972
 A/Title: Crystal structure of human carbonic anhydrase C.
 A/Reference number: A93404; MUID:72111787; PMID:4621826
 A/Contents: annotation; X-ray crystallography, 2.0 angstroms
 A/Note: other residues at the active site are His-64, Asn-67, Tyr-127, Leu-197, Thr-1
 C/Genetics:
 A/Gene: GDB:CA2
 A/Cross-references: GDB:119739; OMIM:259730
 A/Map position: 8q13-qq22.1
 A/Introns: 12/1, 78/1, 117/3, 148/3, 169/3, 221/3
 C/Function:

A:Description: catalyzes the reversible dissociation of carbonic acid to carbon dioxide
 A:Note: this form is expressed in erythrocytes and other tissues; deficiency of this form
 C:Superfamily: carbonate dehydratase, carbonic anhydrase homology
 C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; metalloprotein; mono
 F:2-260/Product: carbonate dehydratase II #status experimental <MAT>
 F:5-250/Domain: carbonic anhydrase homology <CAH>
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental
 F:94,96,119/Binding site: zinc (His) #status experimental

Query Match 15.7%; Score 379.5; DB 1; Length 260;
 Best Local Similarity 34.5%; Pred. No. 3.3e-17;
 Matches 90; Conservative 42; Mismatches 114; Indels 15; Gaps 7;

QY 140 HMRYG---GDPMPRVSPACAGRFQSPVDIRPOLAFCALRPPELLGQLPLPELR 196
 DB 4 HMGYGKNGPBEHMHKDFIAKGERQSPVIDHTAKYDPSLKPSV---SYDAQTSRL 60
 QY 197 NNGHSVQLTPGLEMAL---GP-GREYRALQLHMGAGRPGESEHTVEGRFPALIV 252
 DB 61 NNGHAFNVEFDDSQKAVLKSGPDLGTYYRLIQFHFHMGSLDQGSSEHTVKKKVADELHL 120
 QY 253 VHLSTAFARVDEALGRPGGLAVLAFLIEGEPENSAVEQLSRLEIEAEGSETOVPGLD 312
 DB 121 VHMNTKYGDGFKAVAGQPDGLAVLGIFLKVG--SAKPGQKQVVDVDSIKTKGKADFTNFD 199
 QY 313 ISALLPDSFRFYQEGSLTTPCAQGVITVFNQTMLSAKQL---HTLSDTLWGPDS 369
 DB 180 PGLLPESLD-YMTYPSGLTTPPLECTVTLKEPISVSEQVLKFRKLNFNGEGEFP 238
 QY 370 RLQNFRAOTPLNGRVIEASF 390
 DB 239 LNVDMWRPAQPLKRGVYKASF 259

RESULT 12

CRHB2

carbonate dehydratase (EC 4.2.1.1) II - sheep (tentative sequence)

N:Alternate names: carbonic anhydrase II

C:Species: Ovis orientalis aries; Ovis ammon aries (domestic sheep)

C>Date: 31-Mar-1981 #sequence_revision 31-Mar-1981 #text_change 31-Mar-2000

C:Accession: A01145

R:Tanais, R.U.; Ferrell, R.E.; Tashian, R.E.

Biochim. Biophys. Acta 371, 534-548, 1974

A:Title: Amino acid sequence of sheep carbonic anhydrase C.

A:Reference number: A90598; PMID:75054988; PMID:4215456

A:Accession: A01145

A:Molecule type: protein

A:Residues: 1-259 <FRAN>

R:Paliet, B.; Gallien, J.M.; Sciaky, M.; Laurent, G.; Charrel, M.

Biochim. Biophys. Acta 576, 290-304, 1979

A:Title: Formes moléculaires multiples de l'anhydrase carbonique erythrocytaire ovine.

A:Reference number: A90628; PMID:79145542; PMID:106895

A:Contents: annotation

A:Note: one minor and three major forms were isolated chromatographically. One of these

C:Superfamily: carbonate dehydratase, carbonic anhydrase homology

C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; zinc

F:4-258/Domain: carbonic anhydrase homology <CAH>

F:1/Modified site: acetylated amino end (Ser) #status experimental

F:93,95,118/Binding site: zinc (His) #status predicted

Query Match 15.6%; Score 379; DB 1; Length 259;
 Best Local Similarity 35.5%; Pred. No. 3.5e-17;
 Matches 93; Conservative 41; Mismatches 112; Indels 16; Gaps 8;

QY 140 HMRYG---GDPMPRVSPACAGRFQSPVDIRPOLAFCALRPPELLGQLPLPELR 196
 DB 3 HMGYGKNGPBEHMHKDFPIADGERSQSPVIDTKAVDPDLKPLALLYEQ---AASRRW 59
 QY 197 NNGHSVQLTPGLEMAL---GP-GREYRALQLHMGAGRPGESEHTVEGRFPALIV 252
 DB 60 NNGHSFNVFEFDDSQKAVLKSGPDLGTYYRLIQFHFHMGSLDQGSSEHTVKKKVADELHL 119
 QY 253 VHLSTAFARVDEALGRPGGLAVLAFLIEGEPENSAVEQLSRLEIEAEGSETOVPGLD 312

DB 120 VHMNTKYGDGFKAVAGQPDGLAVLGIFLKVG--DANPALQKLVLDVDSIKTKGKADFPND 178
 QY 313 ISALLPDSFRFYQEGSLTTPCAQGVITVFNQTMLSAKQL---HTLSDTLWGPDS 369
 DB 179 PSSILKRALN-YMTYPSGLTTPPLECTVTLKEPISVSEQVLKFRSLNFVAEGEP 237
 QY 370 RLQNFRAOTPLNGRVIEASF 391
 DB 238 LNVDMWRPAQPLKRGVYKASF 258

RESULT 13

CRHB2

carbonate dehydratase (EC 4.2.1.1) II - rabbit (tentative sequence)

N:Alternate names: carbonic anhydrase II

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 31-Mar-1981 #sequence_revision 31-Mar-1981 #text_change 31-Mar-2000

C:Accession: A01142

R:Ferrell, R.E.; Stroup, S.R.; Tanais, R.J.; Tashian, R.E.

Biochim. Biophys. Acta 533, 1-11, 1978

A:Title: Amino acid sequence of rabbit carbonic anhydrase II.

A:Reference number: A01142; PMID:78144871; PMID:416851

A:Accession: A01142

A:Molecule type: protein

A:Residues: 1-259 <FRB>

A:Note: 203-Glu was also found

C:Superfamily: carbonate dehydratase, carbonic anhydrase homology

C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; zinc

F:4-258/Domain: carbonic anhydrase homology <CAH>

F:1/Modified site: acetylated amino end (Ser) #status experimental

F:93,95,118/Binding site: zinc (His) #status predicted

Query Match 15.4%; Score 374.5; DB 1; Length 259;
 Best Local Similarity 34.5%; Pred. No. 6.8e-17;
 Matches 90; Conservative 42; Mismatches 114; Indels 15; Gaps 8;

QY 140 HMRYG---GDPMPRVSPACAGRFQSPVDIRPOLAFCALRPPELLGQLPLPELR 196
 DB 3 HMGYGKNGPBEHMHKDFPIADGERSQSPVIDTKAVDPDLKPLALLYEQ---ISRLI 59
 QY 197 NNGHSVQLTPGLEMAL---GP-GREYRALQLHMGAGRPGESEHTVEGRFPALIV 252
 DB 60 NNGHSFNVFEFDDSHKDKVLEKGPLEGYYRLIQFHFHMGSLDQGSSEHTVKKKVADELHL 119
 QY 253 VHLSTAFARVDEALGRPGGLAVLAFLIEGEPENSAVEQLSRLEIEAEGSETOVPGLD 312
 DB 120 VHMNTKYGDGFKAVAGQPDGLAVLGIFLKVG--SATPGLQKVVDTLSIKTKGKSVDFNFD 178
 QY 313 ISALLPDSFRFYQEGSLTTPCAQGVITVFNQTMLSAKQL---HTLSDTLWGPDS 369
 DB 179 PGLLPESLD-YMTYPSGLTTPPLECTVTLKEPISVSEQVLKFRSLNFVAEGEP 237
 QY 370 RLQNFRAOTPLNGRVIEASF 390
 DB 238 PNVDMWRPAQPLKRGVYKASF 258

RESULT 14

JH0527

carbonate dehydratase (EC 4.2.1.1) II - rat

N:Alternate names: carbonic anhydrase II

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Jun-1999

C:Accession: JH0527; S20114

R:Stoller, C.A.; McGowan, M.H.; Heim, R.A.; Varia, M.; Neubauer, J.A.

Gene 109, 265-267, 1991

A:Title: Nucleotide sequence of a cDNA encoding rat brain carbonic anhydrase II and

A:Reference number: JH0527; PMID:92112053; PMID:1765271

A:Accession: JH0527

A:Molecule type: mRNA

A:Residues: 1-260 <STO>

A:Cross-references: EMBL:X58294; NID:g55837; PID:CAA41227.1; PID:g55838

A:Experimental source: brain
C:Comment: Carbonate dehydratase is a monomeric zinc metalloenzyme that catalyzes the re
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
C:Keywords: carbon-oxygen lyase; hydro-lyase; zinc
F:5-259/Domain: carbonic anhydrase homology <CAH>

Query Match 15.2%; Score 366.5; DB 2; Length 260;
Best Local Similarity 34.1%; Pred. No. 1,7e-16;
Matches 89; Conservative 41; Mismatches 116; Indels 15; Gaps 7;

QY 140 HWRV--GSDPWRVSPACAGRFQSPVDIRPQLAFCPALPRLLELGFQLPPLRLR 196
DB 4 HMGYSKNGPENHMKERPIANGDROSPVIDTGTQHDPSLQPLLIC--YDKVASKSIV 60
QY 197 NNGHSVQTLTPGLEMAL--GP-GREYRALQHLHWGAAGRPGSEHTVEGHRPPAEIHV 252
DB 61 NNGHSFVNEFDSDQDFVAVKEGFLSGSYRLIQPHFWGSSDQGSSEHTVKKKYAEHL 120
QY 253 VHSSTAFARVDEALGRPGGLAVLAFLERGPENSAVEQLRLLEIAEGSETQVPGLD 312
DB 121 VHWNTKYGDFGKAVQHPDGLAVLGIFLKI-GP-ASQGLKITEALHSIKTKGRAPAFANFD 179
QY 313 ISALPDSFRRFYQYEGSLTTPCAQGVITVFNQVWMLSAKQL--HTLSPTLMGPGDS 369
DB 180 PCSLLPGLND-VWYTPGSLTTPPLLECVTVLKEPTIVSSQMSHFRKLNFSGEAE 238
QY 370 RLQINFRATQPLNGRVLEASF 390
DB 239 LWDVNMWRPAQLNKRKIKASF 259

RESULT 15

CRMOIR
carbonate dehydratase (EC 4.2.1.1) I - rhesus macaque (tentative sequence)
N:Alternate names: carbonic anhydrase I
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 31-Mar-2000
C/Accession: A01139
R/Henriksson, D.; Tanis, R.J.; Taehlan, R.E.
Biochem. Biophys. Res. Commun. 96, 135-142, 1980
A/Title: The amino acid sequence of carbonic anhydrase I from the Rhesus macaque.
A/Reference number: A01139; NID:81062409; PMID:6776950
A/Accession: A01139
A:Molecule type: protein
A:Residues: 1-260 <RHN>
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; zinc
F:5-260/Domain: carbonic anhydrase homology <CAH>
F:1/Modified site: acetylated amino end (Ala) #status experimental
F:94,96,119/Binding site: zinc (His) #status predicted

Query Match 15.1%; Score 366.5; DB 1; Length 260;
Best Local Similarity 32.6%; Pred. No. 2.2e-16;
Matches 87; Conservative 48; Mismatches 109; Indels 23; Gaps 8;

QY 132 DKRGDDQSHWRVYGDPPWPRVSPACAGRFQSPVDIRPQLAFCPALPRLLELGFQLPPLP 191
DB 9 DKRGPEQ-----MSKLYPIANGNNQSPVDIKTSEAKHDTSLKPLISV---SYNPAT 55
QY 192 ELKRNNGHGVQTLTPGLEMAL--GP-GREYRALQHLHWGAAGRPGSEHTVEGHRFP 247
DB 56 AKELINVGSHFHNFDNDRSVLKGPFSDSRLFOFHFWGSSNEVGSSEHTVGVVYS 115
QY 248 AEIHVVALSTA-PARVDEALGRPGGLAVLAFLERGPENSAVEQLRLLEIAEGSET 306
DB 116 SELHIYHMSAKYSLSLAENASKADGLAVIGVLMKVG-EANPKLQVLDALHAIKTKGRRA 174
QY 307 QVPGDLISALLPDSFRRFYQYEGSLTTPCAQGVITVFNQVWMLSAKQL--HTLSPTL 363
DB 175 PFTNFPSTLLPGLD-FWYTPGSLTTPPLLECVTVLKEPTIVSSQMSHFRKLNFSLSNV 233
QY 364 WGPDSRLQINFRATQPLNGRVLEASF 390

DB 234 EGSNPFVIOANNRPTQPLKGRIVRAS 260
Search completed: March 5, 2004, 09:07:53
Job time: 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2004, 09:03:00 ; Search time 22 seconds
(without alignments)
1077.105 Million cell updates/sec

Title: US-09-967-237-2

Perfect score: 2424
Sequence: 1 MAPLCSPWLPILIPAPAPG.....RRGKGVSVPAEVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6C_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2424	100.0	459	2	US-08-481-658B-2
2	2424	100.0	459	2	US-08-477-504A-2
3	2424	100.0	459	2	US-08-486-756A-2
4	2424	100.0	459	2	US-08-485-862B-2
5	2424	100.0	459	2	US-08-787-739-2
6	2424	100.0	459	3	US-08-487-077A-2
7	2424	100.0	459	3	US-08-485-863A-2
8	2424	100.0	459	3	US-08-485-049D-2
9	2424	100.0	459	3	US-09-178-115-2
10	2424	100.0	459	3	US-09-177-776-2
11	2227	91.9	422	3	US-08-335-469-2
12	2007	82.8	377	3	US-08-787-739-87
13	2007	82.8	377	3	US-09-178-115-87
14	2007	82.8	377	3	US-09-177-776-87
15	1398.5	57.7	429	1	US-07-664-588-2
16	1398.5	57.7	429	5	PCT-US83-02024-2
17	1370	56.5	237	3	US-08-787-739-51
18	1370	56.5	237	3	US-09-178-115-51
19	1370	56.5	257	3	US-09-177-776-51
20	1364	56.3	256	2	US-08-481-658B-51
21	1364	56.3	256	2	US-08-477-504A-51
22	1364	56.3	256	2	US-08-486-756A-51
23	1364	56.3	256	2	US-08-485-862B-51
24	1364	56.3	256	2	US-08-487-077A-51
25	1364	56.3	256	3	US-08-485-863A-51
26	1364	56.3	256	3	US-08-485-049D-51
27	884	36.5	170	2	US-08-481-658B-54

28	884	36.5	170	2	US-08-477-504A-54	Sequence 54, Appl
29	884	36.5	170	2	US-08-486-756A-54	Sequence 54, Appl
30	884	36.5	170	2	US-08-485-862B-54	Sequence 54, Appl
31	884	36.5	170	3	US-08-487-077A-54	Sequence 54, Appl
32	884	36.5	170	3	US-08-485-863A-54	Sequence 54, Appl
33	884	36.5	170	3	US-08-485-049D-54	Sequence 54, Appl
34	537.5	22.2	354	1	US-08-276-919-2	Sequence 2, Appl
35	537.5	22.2	354	1	US-08-776-088-2	Sequence 2, Appl
36	537.5	22.2	354	1	US-08-776-088-6	Sequence 2, Appl
37	537.5	22.2	354	3	US-09-325-320-2	Sequence 2, Appl
38	537.5	22.2	354	3	US-09-585-109-2	Sequence 2, Appl
39	537.5	22.2	354	3	PCT-US95-09145A-2	Sequence 2, Appl
40	537.5	22.2	354	5	PCT-US95-09145A-6	Sequence 6, Appl
41	535	22.1	325	1	US-08-276-919-4	Sequence 4, Appl
42	535	22.1	325	1	US-08-776-088-4	Sequence 4, Appl
43	535	22.1	325	5	PCT-US95-09145A-4	Sequence 4, Appl
44	533	22.0	325	1	US-08-276-919-13	Sequence 13, Appl
45	533	22.0	325	1	US-08-776-088-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-2
Sequence 2, Application US/08481658B
Patent No. 5955075
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: Mn Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: First 37 amino acids represent
signal peptide, and remaining amino acids
DESCRIPTION: represent mature protein
US-08-481-658B-2
Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 8,6e-197;

Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLIPAPAPGLTVQLLSLLLMVPHQRLPRMOEDSPUGSSSGEDDL 60
DB 1 MAPLCSPWMLPLIPAPAPGLTVQLLSLLLMVPHQRLPRMOEDSPUGSSSGEDDL 60

QY 61 GCEDLPSSEDSPREDDPGEBDLPGEBDLPGEBDLPGEBDLPGEBDLPGEBDL 120
DB 61 GCEDLPSSEDSPREDDPGEBDLPGEBDLPGEBDLPGEBDLPGEBDLPGEBDL 120

QY 121 DPOEPQNNARHDKEDDQSHWRYGDDPPMPRVSPACAGFQSPVDIRPQLAFCALRPL 180
DB 121 DPOEPQNNARHDKEDDQSHWRYGDDPPMPRVSPACAGFQSPVDIRPQLAFCALRPL 180

QY 181 ELIGFQLPPLPELRRLNNGHSVQLTLPGLMALGPRGRYRALQLHLHGAAGRPSGSH 240
DB 181 ELIGFQLPPLPELRRLNNGHSVQLTLPGLMALGPRGRYRALQLHLHGAAGRPSGSH 240

QY 241 VEGHFPFAEIHVHLSLTAFAVDEALGPRGLAVLAFLBEGPEENSAYEQLSLLEIA 300
DB 241 VEGHFPFAEIHVHLSLTAFAVDEALGPRGLAVLAFLBEGPEENSAYEQLSLLEIA 300

QY 301 BEGSETQVPGDLISALLPSDFSRYFYEGSLTTPPCAQGVITVFNQVWLSAKQHLTIS 360
DB 301 BEGSETQVPGDLISALLPSDFSRYFYEGSLTTPPCAQGVITVFNQVWLSAKQHLTIS 360

QY 361 DTLMGPGDSRLQNLFRATQPLNGRVITASFPAGVDSPPRAAPVQVINSCLAADILALVF 420
DB 361 DTLMGPGDSRLQNLFRATQPLNGRVITASFPAGVDSPPRAAPVQVINSCLAADILALVF 420

QY 421 GILFAVTSVAFVLQMRQRHRRGTGKGVSYRPAEVAETGA 459
DB 421 GILFAVTSVAFVLQMRQRHRRGTGKGVSYRPAEVAETGA 459

RESULT 2
US-08-477-504A-2
Sequence 2, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: First 37 amino acids represent
DESCRIPTION: signal peptide, and remaining amino acids
DESCRIPTION: represent mature protein

US-08-477-504A-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 8.6e-197;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLIPAPAPGLTVQLLSLLLMVPHQRLPRMOEDSPUGSSSGEDDL 60
DB 1 MAPLCSPWMLPLIPAPAPGLTVQLLSLLLMVPHQRLPRMOEDSPUGSSSGEDDL 60

QY 61 GCEDLPSSEDSPREDDPGEBDLPGEBDLPGEBDLPGEBDLPGEBDLPGEBDL 120
DB 61 GCEDLPSSEDSPREDDPGEBDLPGEBDLPGEBDLPGEBDLPGEBDLPGEBDL 120

QY 121 DPOEPQNNARHDKEDDQSHWRYGDDPPMPRVSPACAGFQSPVDIRPQLAFCALRPL 180
DB 121 DPOEPQNNARHDKEDDQSHWRYGDDPPMPRVSPACAGFQSPVDIRPQLAFCALRPL 180

QY 181 ELIGFQLPPLPELRRLNNGHSVQLTLPGLMALGPRGRYRALQLHLHGAAGRPSGSH 240
DB 181 ELIGFQLPPLPELRRLNNGHSVQLTLPGLMALGPRGRYRALQLHLHGAAGRPSGSH 240

QY 241 VEGHFPFAEIHVHLSLTAFAVDEALGPRGLAVLAFLBEGPEENSAYEQLSLLEIA 300
DB 241 VEGHFPFAEIHVHLSLTAFAVDEALGPRGLAVLAFLBEGPEENSAYEQLSLLEIA 300

QY 301 BEGSETQVPGDLISALLPSDFSRYFYEGSLTTPPCAQGVITVFNQVWLSAKQHLTIS 360
DB 301 BEGSETQVPGDLISALLPSDFSRYFYEGSLTTPPCAQGVITVFNQVWLSAKQHLTIS 360

QY 361 DTLMGPGDSRLQNLFRATQPLNGRVITASFPAGVDSPPRAAPVQVINSCLAADILALVF 420
DB 361 DTLMGPGDSRLQNLFRATQPLNGRVITASFPAGVDSPPRAAPVQVINSCLAADILALVF 420

QY 421 GILFAVTSVAFVLQMRQRHRRGTGKGVSYRPAEVAETGA 459
DB 421 GILFAVTSVAFVLQMRQRHRRGTGKGVSYRPAEVAETGA 459

RESULT 3
US-08-486-756A-2
Sequence 2, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/260,190
 FILING DATE: 15-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Lauder, Leona L.
 REGISTRATION NUMBER: 30,863
 REFERENCE/DOCKET NUMBER: D-0021.3C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-435-2034
 TELEFAX: 415-435-0727
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 459 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 DESCRIPTION: First 37 amino acids represent
 signal peptide, and remaining amino acids
 represent mature protein
 US-08-486-756A-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
 Best Local Similarity 100.0%; Pred. No. 8.6e-197;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSMWLPLIIPAPAGLTIVQLLSLLLMVHPQLPRMOWDSPLGGSSGEDDPL 60
 DB 1 MAPLCPSMWLPLIIPAPAGLTIVQLLSLLLMVHPQLPRMOWDSPLGGSSGEDDPL 60
 QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPQEPONNAHRDKEDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCFALRPL 180
 DB 121 DPQEPONNAHRDKEDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCFALRPL 180
 QY 181 ELIGFQLPPLPELRIRNNHSHVOLTPPGLEMALGPRERVALQLHLMWGAAGRPGSEHT 240
 DB 181 ELIGFQLPPLPELRIRNNHSHVOLTPPGLEMALGPRERVALQLHLMWGAAGRPGSEHT 240
 QY 241 VEGHRFPALIHVHLSTAFARVDEALGRPGGLAVLAFLFEEGPEENSAYEQLLSRLEIEA 300
 DB 241 VEGHRFPALIHVHLSTAFARVDEALGRPGGLAVLAFLFEEGPEENSAYEQLLSRLEIEA 300
 QY 301 EEGSETQVPGGLDISALLPSDFSRFYQYEGSLTPPCAQGVITWVFNQTMLSAKQHLTIS 360
 DB 301 EEGSETQVPGGLDISALLPSDFSRFYQYEGSLTPPCAQGVITWVFNQTMLSAKQHLTIS 360
 QY 361 DTLMGGBSRQLQNFATQPLNGRVTEASFPAGVDSPPRAAPVQLNSCLAAGDILLVLF 420
 DB 361 DTLMGGBSRQLQNFATQPLNGRVTEASFPAGVDSPPRAAPVQLNSCLAAGDILLVLF 420
 QY 421 GLFPAVTSVAFVQMERQHRRTGKGVSYRPAEVAETGA 459
 DB 421 GLFPAVTSVAFVQMERQHRRTGKGVSYRPAEVAETGA 459

RESULT 4
 US-08-485-862B-2
 Sequence 2, Application US/08485862B
 Patent No. 5989838
 GENERAL INFORMATION:
 APPLICANT: Zavada, Jan
 APPLICANT: Paetorekova, Silvia
 APPLICANT: Paetorek, Jaromir
 TITLE OF INVENTION: MN Gene and Protein
 NUMBER OF SEQUENCES: 86
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Leona L. Lauder
 STREET: 6 Mariposa Court
 CITY: Tiburon

STATE: California
 COUNTRY: USA
 ZIP: 94920
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,862B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/477,504
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/260,190
 FILING DATE: 15-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Lauder, Leona L.
 REGISTRATION NUMBER: 30,863
 REFERENCE/DOCKET NUMBER: D-0021.3D
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-435-2034
 TELEFAX: 415-435-0727
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 459 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 DESCRIPTION: First 37 amino acids represent
 signal peptide, and remaining amino acids
 represent mature protein
 US-08-485-862B-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
 Best Local Similarity 100.0%; Pred. No. 8.6e-197;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSMWLPLIIPAPAGLTIVQLLSLLLMVHPQLPRMOWDSPLGGSSGEDDPL 60
 DB 1 MAPLCPSMWLPLIIPAPAGLTIVQLLSLLLMVHPQLPRMOWDSPLGGSSGEDDPL 60
 QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPQEPONNAHRDKEDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCFALRPL 180
 DB 121 DPQEPONNAHRDKEDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPOLAFCFALRPL 180
 QY 181 ELIGFQLPPLPELRIRNNHSHVOLTPPGLEMALGPRERVALQLHLMWGAAGRPGSEHT 240
 DB 181 ELIGFQLPPLPELRIRNNHSHVOLTPPGLEMALGPRERVALQLHLMWGAAGRPGSEHT 240
 QY 241 VEGHRFPALIHVHLSTAFARVDEALGRPGGLAVLAFLFEEGPEENSAYEQLLSRLEIEA 300
 DB 241 VEGHRFPALIHVHLSTAFARVDEALGRPGGLAVLAFLFEEGPEENSAYEQLLSRLEIEA 300
 QY 301 EEGSETQVPGGLDISALLPSDFSRFYQYEGSLTPPCAQGVITWVFNQTMLSAKQHLTIS 360
 DB 301 EEGSETQVPGGLDISALLPSDFSRFYQYEGSLTPPCAQGVITWVFNQTMLSAKQHLTIS 360
 QY 361 DTLMGGBSRQLQNFATQPLNGRVTEASFPAGVDSPPRAAPVQLNSCLAAGDILLVLF 420
 DB 361 DTLMGGBSRQLQNFATQPLNGRVTEASFPAGVDSPPRAAPVQLNSCLAAGDILLVLF 420
 QY 421 GLFPAVTSVAFVQMERQHRRTGKGVSYRPAEVAETGA 459
 DB 421 GLFPAVTSVAFVQMERQHRRTGKGVSYRPAEVAETGA 459

RESULT 5

US-08-787-739-2
Sequence 2, Application US/08787739
Patent No. 6027887
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jarmir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
DESCRIPTION: First 37 amino acids represent
DESCRIPTION: signal peptide, and remaining amino acids
DESCRIPTION: represent mature protein
US-08-787-739-2

Query Match 100.0%; Score 2424; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 8.6e-197;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSFWPLFLIPAPAGITVQLLSLLLMFVHPORLPRMOEDSPICGGSSGSDPL 60
DB 1 MAPLCPSFWPLFLIPAPAGITVQLLSLLLMFVHPORLPRMOEDSPICGGSSGSDPL 60
QY 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPVVKPSEEGSLKLEDLPTVEAPG 120

DB 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPVVKPSEEGSLKLEDLPTVEAPG 120
QY 121 DPOEPQNNARDEKGDQSHWRYGDPMPWRVSPACAGRFQSPVDIRPOLAFCALRPL 180
DB 121 DPOEPQNNARDEKGDQSHWRYGDPMPWRVSPACAGRFQSPVDIRPOLAFCALRPL 180
QY 181 ELIGFQLPPLPELRNNHSHVQLTPCLEMALPGREYRLQLHLMGAAGRGSEHT 240
DB 181 ELIGFQLPPLPELRNNHSHVQLTPCLEMALPGREYRLQLHLMGAAGRGSEHT 240
QY 241 VEGHPPAEIHVHLSAPARVDEALGRGGLAVLAFLBEPSEENSAVEQLSRLEIA 300
DB 241 VEGHPPAEIHVHLSAPARVDEALGRGGLAVLAFLBEPSEENSAVEQLSRLEIA 300
QY 301 EEGSETQVPLDLSALLBEDFSRYFEGSLTTPPCAGQVITVNTQVIMSAKQLHTLS 360
DB 301 EEGSETQVPLDLSALLBEDFSRYFEGSLTTPPCAGQVITVNTQVIMSAKQLHTLS 360
QY 361 DTLMGPDSSRLQINFRATQPLNGRVTEASFPAQVDSPPRAEPVOLNSCLAAGDILALVF 420
DB 361 DTLMGPDSSRLQINFRATQPLNGRVTEASFPAQVDSPPRAEPVOLNSCLAAGDILALVF 420
QY 421 GLIFATSVAFVQMRQRHRTKGVSYRPAEVAETGA 459
DB 421 GLIFATSVAFVQMRQRHRTKGVSYRPAEVAETGA 459

RESULT 6

US-08-487-077A-2
Sequence 2, Application US/08487077A
Patent No. 6069242

GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jarmir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,077A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3H
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: First 37 amino acids represent
DESCRIPTION: signal peptide, and remaining amino acids

DESCRIPTION: represent mature protein
US-08-487-077A-2

Query Match 100.0%; Score 2424; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 8,6e-197;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWLPILIPAPAGLTQVLLSLILMPVHPQRLPRMOEDSPLOGSSGSDPL 60
DB 1 MAPLCSPWLPILIPAPAGLTQVLLSLILMPVHPQRLPRMOEDSPLOGSSGSDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPEPQNNAHKDEKDDQSHWRVYGGDPMPWRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPEPQNNAHKDEKDDQSHWRVYGGDPMPWRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELIGFQLPPLPELRIRNNHSHVQLTPPGLEMALGPGEYRALQHLHWGAAGRGSEHT 240
DB 181 ELIGFQLPPLPELRIRNNHSHVQLTPPGLEMALGPGEYRALQHLHWGAAGRGSEHT 240
QY 241 VEGHRRPFAIHHVHSTAFARVDALGRPGGLAVLAFLBEGPEENSAVEQLLSRLEIA 300
DB 241 VEGHRRPFAIHHVHSTAFARVDALGRPGGLAVLAFLBEGPEENSAVEQLLSRLEIA 300
QY 301 EBSSETQVPGDLISALLPSDSFRYFQYEGSLTTPPCAQGVITWTFNQVWLSAKQHTLS 360
DB 301 EBSSETQVPGDLISALLPSDSFRYFQYEGSLTTPPCAQGVITWTFNQVWLSAKQHTLS 360
QY 361 DTLMGPGDSRLQINFRATQPLNGRVIEASFPAGVSSPRAEPVQVNSCLAAGDILALVF 420
DB 361 DTLMGPGDSRLQINFRATQPLNGRVIEASFPAGVSSPRAEPVQVNSCLAAGDILALVF 420
QY 421 GLIFAVTSVAFVQMRQRHRRGTGKGVSYRAEVAETGA 459
DB 421 GLIFAVTSVAFVQMRQRHRRGTGKGVSYRAEVAETGA 459

RESULT 7

US-08-485-863A-2
Sequence 2, Application US/08485863A
Patent No. 6093548
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPC)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,863A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3G

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: protein

DESCRIPTION: First 37 amino acids represent

DESCRIPTION: signal peptide, and remaining amino acids

US-08-485-863A-2

Query Match 100.0%; Score 2424; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 8,6e-197;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWLPILIPAPAGLTQVLLSLILMPVHPQRLPRMOEDSPLOGSSGSDPL 60
DB 1 MAPLCSPWLPILIPAPAGLTQVLLSLILMPVHPQRLPRMOEDSPLOGSSGSDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPEPQNNAHKDEKDDQSHWRVYGGDPMPWRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPEPQNNAHKDEKDDQSHWRVYGGDPMPWRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELIGFQLPPLPELRIRNNHSHVQLTPPGLEMALGPGEYRALQHLHWGAAGRGSEHT 240
DB 181 ELIGFQLPPLPELRIRNNHSHVQLTPPGLEMALGPGEYRALQHLHWGAAGRGSEHT 240
QY 241 VEGHRRPFAIHHVHSTAFARVDALGRPGGLAVLAFLBEGPEENSAVEQLLSRLEIA 300
DB 241 VEGHRRPFAIHHVHSTAFARVDALGRPGGLAVLAFLBEGPEENSAVEQLLSRLEIA 300
QY 301 EBSSETQVPGDLISALLPSDSFRYFQYEGSLTTPPCAQGVITWTFNQVWLSAKQHTLS 360
DB 301 EBSSETQVPGDLISALLPSDSFRYFQYEGSLTTPPCAQGVITWTFNQVWLSAKQHTLS 360
QY 361 DTLMGPGDSRLQINFRATQPLNGRVIEASFPAGVSSPRAEPVQVNSCLAAGDILALVF 420
DB 361 DTLMGPGDSRLQINFRATQPLNGRVIEASFPAGVSSPRAEPVQVNSCLAAGDILALVF 420
QY 421 GLIFAVTSVAFVQMRQRHRRGTGKGVSYRAEVAETGA 459
DB 421 GLIFAVTSVAFVQMRQRHRRGTGKGVSYRAEVAETGA 459

RESULT 8

US-08-485-049D-2
Sequence 2, Application US/08485049D
Patent No. 6204370
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,049D
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: First 37 amino acids represent
signal peptide, and remaining amino acids
DESCRIPTION: represent mature protein
US-08-485-049D-2

Query Match 100.0%; Score 2424; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 8.6e-197;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLIPAPAPGLTVQLLSLILMPVHPORLPFMOEDSPFGGSSGSDPPL 60
DB 1 MAPLCSPWMLPLIPAPAPGLTVQLLSLILMPVHPORLPFMOEDSPFGGSSGSDPPL 60
QY 61 GSEDLPSSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GSEDLPSSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPOBPNNAHREKGGDQSHWRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPOBPNNAHREKGGDQSHWRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELLGFQPLPELRLNNGSHSVQTLPGLEMALGPREYRALQHLHMGAGRPGSEHT 240
DB 181 ELLGFQPLPELRLNNGSHSVQTLPGLEMALGPREYRALQHLHMGAGRPGSEHT 240
QY 241 VEGHRFPALHVVHLSFAFARVDEALGRPGGLAVLAFLBEGPEENSAYEQLLSRLEBIA 300
DB 241 VEGHRFPALHVVHLSFAFARVDEALGRPGGLAVLAFLBEGPEENSAYEQLLSRLEBIA 300
QY 301 EEGSETQVPGDLISALLPSDFSRFYQEGSLTTPCAQGVITVFNQTVMLSAKQHLTSL 360
DB 301 EEGSETQVPGDLISALLPSDFSRFYQEGSLTTPCAQGVITVFNQTVMLSAKQHLTSL 360
QY 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAVDSSPRAAPVQINSCLAAGDILALVF 420
DB 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAVDSSPRAAPVQINSCLAAGDILALVF 420
QY 421 GLLPATVSVAFLVQMRQRGRGTGKGVSYRPAEVAETGA 459
DB 421 GLLPATVSVAFLVQMRQRGRGTGKGVSYRPAEVAETGA 459

RESULT 9
US-09-178-115-2
Sequence 2, Application US/09178115
Patent No. 6297041
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein

FILE REFERENCE: D-0021.5A
CURRENT APPLICATION NUMBER: US/09/178,115
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 09/177,776
EARLIER FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 08/787,739
EARLIER FILING DATE: 1997-01-24
EARLIER APPLICATION NUMBER: 08/485,049
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/486,756
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/477,504
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/481,658
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,862
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,863
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/487,077
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/260,190
EARLIER FILING DATE: 1994-06-15
EARLIER APPLICATION NUMBER: 08/177,093
EARLIER FILING DATE: 1993-12-30
EARLIER APPLICATION NUMBER: 07/964,589
EARLIER FILING DATE: 1992-10-21
EARLIER APPLICATION NUMBER: PV-709-92
EARLIER FILING DATE: 1992-03-11
NUMBER OF SEQ ID NOS: 116
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 459
TYPE: PRT
ORGANISM: HUMAN
US-09-178-115-2

Query Match 100.0%; Score 2424; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 8.6e-197;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLIPAPAPGLTVQLLSLILMPVHPORLPFMOEDSPFGGSSGSDPPL 60
DB 1 MAPLCSPWMLPLIPAPAPGLTVQLLSLILMPVHPORLPFMOEDSPFGGSSGSDPPL 60
QY 61 GSEDLPSSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GSEDLPSSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPOBPNNAHREKGGDQSHWRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPOBPNNAHREKGGDQSHWRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELLGFQPLPELRLNNGSHSVQTLPGLEMALGPREYRALQHLHMGAGRPGSEHT 240
DB 181 ELLGFQPLPELRLNNGSHSVQTLPGLEMALGPREYRALQHLHMGAGRPGSEHT 240
QY 241 VEGHRFPALHVVHLSFAFARVDEALGRPGGLAVLAFLBEGPEENSAYEQLLSRLEBIA 300
DB 241 VEGHRFPALHVVHLSFAFARVDEALGRPGGLAVLAFLBEGPEENSAYEQLLSRLEBIA 300
QY 301 EEGSETQVPGDLISALLPSDFSRFYQEGSLTTPCAQGVITVFNQTVMLSAKQHLTSL 360
DB 301 EEGSETQVPGDLISALLPSDFSRFYQEGSLTTPCAQGVITVFNQTVMLSAKQHLTSL 360
QY 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAVDSSPRAAPVQINSCLAAGDILALVF 420
DB 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAVDSSPRAAPVQINSCLAAGDILALVF 420
QY 421 GLLPATVSVAFLVQMRQRGRGTGKGVSYRPAEVAETGA 459
DB 421 GLLPATVSVAFLVQMRQRGRGTGKGVSYRPAEVAETGA 459

RESULT 10

US-09-177-776-2

Sequence 2, Application US/09177776A

Patent No. 6237051

GENERAL INFORMATION:

APPLICANT: Zavada, Jan

APPLICANT: Pastorekova, Silvia

APPLICANT: Pastorek, Jaromir

TITLE OF INVENTION: MN Gene and Protein

FILE REFERENCE: D-0021 SA

CURRENT APPLICATION NUMBER: US/09/177,776A

EARLIER FILING DATE: 1998-10-23

EARLIER APPLICATION NUMBER: 08/787,739

EARLIER FILING DATE: 1997-01-24

EARLIER APPLICATION NUMBER: 08/485,049

EARLIER FILING DATE: 1995-06-07

EARLIER APPLICATION NUMBER: 08/486,756

EARLIER FILING DATE: 1995-06-07

EARLIER APPLICATION NUMBER: 08/477,504

EARLIER FILING DATE: 1995-06-07

EARLIER APPLICATION NUMBER: 08/481,658

EARLIER FILING DATE: 1995-06-07

EARLIER APPLICATION NUMBER: 08/260,190

EARLIER FILING DATE: 1994-06-15

EARLIER APPLICATION NUMBER: 08/177,093

EARLIER FILING DATE: 1993-12-30

EARLIER APPLICATION NUMBER: 07/964,589

EARLIER FILING DATE: 1992-10-21

EARLIER APPLICATION NUMBER: PV-709-92

EARLIER FILING DATE: 1992-03-11

NUMBER OF SEQ ID NOS: 116

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 459

TYPE: PRT

ORGANISM: HUMAN

US-09-177-776-2

Query Match 100.0%; Score 2424; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 8.6e-197; Indels 0; Gaps 0;
Matches 459; Conservative 0; Mismatches 0;

QY 1 MAPLCSPWLPPLIPAPAPAGLTVOQLLSLLIMLVHQRPLPMQEDSPILGGSSGDDPL 60
DB 1 MAPLCSPWLPPLIPAPAPAGLTVOQLLSLLIMLVHQRPLPMQEDSPILGGSSGDDPL 60
QY 61 GSEDLPSREDSREDDPGEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GSEDLPSREDSREDDPGEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPOEPONNAHRKEDDQSHRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPOEPONNAHRKEDDQSHRYGDDPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELIGFOLPPLPRLRNNGHSVQTLTPGLGEMALGPGRETRALQJLHWGAAGPGSEHT 240
DB 181 ELIGFOLPPLPRLRNNGHSVQTLTPGLGEMALGPGRETRALQJLHWGAAGPGSEHT 240
QY 241 VEGHRFPALIHVHLSTAFARVDEALGRPGGLAVLAFAFLEEGPENSAVYEQLSRLBEIA 300
DB 241 VEGHRFPALIHVHLSTAFARVDEALGRPGGLAVLAFAFLEEGPENSAVYEQLSRLBEIA 300
QY 301 EGGSETQVPGDLISALLSSDPGRFYQVHGSILTPPCQGVYIWFVFNQTVMLSAKQJHTLS 360
DB 301 EGGSETQVPGDLISALLSSDPGRFYQVHGSILTPPCQGVYIWFVFNQTVMLSAKQJHTLS 360

RESULT 11

US-08-335-469-2

Sequence 2, Application US/08335469A

Patent No. 6004535

GENERAL INFORMATION:

APPLICANT: Zavada, Jan

APPLICANT: Pastorekova, Silvia

APPLICANT: Pastorek, Jaromir

TITLE OF INVENTION: Methods to Detect and Quantify MN Protein/Polypeptide

FILE REFERENCE: D-0021A

CURRENT APPLICATION NUMBER: US/08/335,469A

EARLIER FILING DATE: 1994-11-07

EARLIER APPLICATION NUMBER: 07/964,589

EARLIER FILING DATE: 1992-10-21

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 422

TYPE: PRT

ORGANISM: Human

US-08-335-469-2

Query Match 91.9%; Score 2227; DB 3; Length 422;
Best Local Similarity 100.0%; Pred. No. 3.5e-180; Indels 0; Gaps 0;
Matches 422; Conservative 0; Mismatches 0;

QY 38 QRLPRMOEDSPILGGSSGDDPLGEEDLPSEEDSPREDDPGEDLPGEEDLPGEEDLP 97
DB 1 QRLPRMOEDSPILGGSSGDDPLGEEDLPSEEDSPREDDPGEDLPGEEDLPGEEDLP 97
QY 98 VKPSEEGSLKLEDDLPVEAPDPOEPONNAHRKEDDQSHRYGDDPMPRVSPACA 157
DB 98 VKPSEEGSLKLEDDLPVEAPDPOEPONNAHRKEDDQSHRYGDDPMPRVSPACA 157
QY 61 VKPSEEGSLKLEDDLPVEAPDPOEPONNAHRKEDDQSHRYGDDPMPRVSPACA 120
DB 61 VKPSEEGSLKLEDDLPVEAPDPOEPONNAHRKEDDQSHRYGDDPMPRVSPACA 120
QY 158 GRFQSPVDIRPOLAFCPALRPLRLRNNGHSVQTLTPGLGEMALGPG 217
DB 121 GRFQSPVDIRPOLAFCPALRPLRLRNNGHSVQTLTPGLGEMALGPG 180
QY 218 REYALQJLHWGAAGPGSEHTVEGHRFPALIHVHLSTAFARVDEALGRPGGLAVLA 277
DB 181 REYALQJLHWGAAGPGSEHTVEGHRFPALIHVHLSTAFARVDEALGRPGGLAVLA 240
QY 278 FLEEGPENSAVYEQLSRLBEIAEGSETQVPGDLISALLSPDSFYQVHGSILTPPCA 337
DB 241 FLEEGPENSAVYEQLSRLBEIAEGSETQVPGDLISALLSPDSFYQVHGSILTPPCA 300
QY 338 QGVYIWFVFNQTVMLSAKQJHTLSLTPGPDSSILQNFATQPLNGRVIEASFPAGVDS 397
DB 301 QGVYIWFVFNQTVMLSAKQJHTLSLTPGPDSSILQNFATQPLNGRVIEASFPAGVDS 360
QY 398 PRAAEVQVNSCLAAGITLAVGGLFAYTSVAFLVQMRORHRGKGVSTRPAVAET 457
DB 361 PRAAEVQVNSCLAAGITLAVGGLFAYTSVAFLVQMRORHRGKGVSTRPAVAET 420
QY 458 GA 459
DB 421 GA 422

RESULT 12

US-08-787-739-87

Sequence 87, Application US/08787739

Patent No. 6027887

GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: MN protein extracellular domain
US-08-787-739-87

Query Match 82.8%; Score 2007; DB 3; Length 377;
Best Local Similarity 100.0%; Pred. No. 1.2e-161;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 QRLPMQDSPLGGSSGSDPLGSEDLPSSEDSPREDDPGEDLPGEEDLPGEEDLPE 97
DB 1 QRLPMQDSPLGGSSGSDPLGSEDLPSSEDSPREDDPGEDLPGEEDLPGEEDLPE 60
QY VKPKSEEGSLKLEDLPTVEAPGDPQEPONNARHDKGDQSHMRYGDDPMPRVSPACA 157
DB 61 VKPKSEEGSLKLEDLPTVEAPGDPQEPONNARHDKGDQSHMRYGDDPMPRVSPACA 120
QY 158 GRFQSPVDIRPOLAFCFALRPILLGFOLPPLBELRLNNGHGSVOLTLPPGLEMALGFG 217
DB 121 GRFQSPVDIRPOLAFCFALRPILLGFOLPPLBELRLNNGHGSVOLTLPPGLEMALGFG 180

QY 218 REFRALQHLHWGAAGRPSSEHTVEGHRPAEIHVHSTARVDEALGRPGGLAVLAA 277
DB 181 REFRALQHLHWGAAGRPSSEHTVEGHRPAEIHVHSTARVDEALGRPGGLAVLAA 240
QY 278 FLEEGPEENSAVEQLSLREELIAEGSETEQVGLDISALLPDEFERYQYEGSLTTPPCA 337
DB 241 FLEEGPEENSAVEQLSLREELIAEGSETEQVGLDISALLPDEFERYQYEGSLTTPPCA 300
QY 338 QGVITWYFNQTVWLSAKQHTLSDTLMGGGDSRLQNRATQPLNGRYTEASFPGVYSS 397
DB 301 QGVITWYFNQTVWLSAKQHTLSDTLMGGGDSRLQNRATQPLNGRYTEASFPGVYSS 360
QY 398 PRAAPYQUNSCAAGD 414
DB 361 PRAAPYQUNSCAAGD 377

RESULT 13
US-09-178-115-87
Sequence 87, Application US/09178115
Patent No. 6297041
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
FILE REFERENCE: D-0021.5A
CURRENT APPLICATION NUMBER: US/09/178,115
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 09/177,776
EARLIER FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 08/787,739
EARLIER FILING DATE: 1997-01-24
EARLIER APPLICATION NUMBER: 08/485,049
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/486,756
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/477,504
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/481,658
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,862
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,863
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/487,077
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/260,190
EARLIER FILING DATE: 1994-06-15
EARLIER APPLICATION NUMBER: 08/177,093
EARLIER FILING DATE: 1993-12-30
EARLIER APPLICATION NUMBER: 07/964,589
EARLIER FILING DATE: 1992-10-21
EARLIER APPLICATION NUMBER: PV-709-92
EARLIER FILING DATE: 1992-03-11
NUMBER OF SEQ ID NOS: 116
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 87
LENGTH: 377
TYPE: PRT
ORGANISM: HUMAN
US-09-178-115-87

Query Match 82.8%; Score 2007; DB 3; Length 377;
Best Local Similarity 100.0%; Pred. No. 1.2e-161;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 QRLPMQDSPLGGSSGSDPLGSEDLPSSEDSPREDDPGEDLPGEEDLPGEEDLPE 97
DB 1 QRLPMQDSPLGGSSGSDPLGSEDLPSSEDSPREDDPGEDLPGEEDLPGEEDLPE 60
QY VKPKSEEGSLKLEDLPTVEAPGDDPQEPONNARHDKGDQSHMRYGDDPMPRVSPACA 157
DB 98 VKPKSEEGSLKLEDLPTVEAPGDDPQEPONNARHDKGDQSHMRYGDDPMPRVSPACA 157

Db 61 VKKSEEGSLKLEDPVTEARQDPEQNNAHKXGDDQSHNRVGGDPPMPVSPACA 120
Qy 158 GRQSPVDIRPOLAACCPALRPELIGFQLPELRIRNNGHSVQLTPPGLEMALGPG 217
Db 121 GRQSPVDIRPOLAACCPALRPELIGFQLPELRIRNNGHSVQLTPPGLEMALGPG 180
Qy 218 REYRALQHLHMGAGRGSEHTVEGRFPFAIHVHLSTAFARVDALGRPGGLAVLAA 277
Db 181 REYRALQHLHMGAGRGSEHTVEGRFPFAIHVHLSTAFARVDALGRPGGLAVLAA 240
Qy 278 FLEEGPENSAYEQLSLRLEIAEGSETQVPGDLISALPDSFRYQYEGSLTTPCA 337
Db 241 FLEEGPENSAYEQLSLRLEIAEGSETQVPGDLISALPDSFRYQYEGSLTTPCA 300
Qy 338 QGVITVFNQVTWLSAKQHTLSDTLWPGDSRLQINFRATQPLNGRVITASFPAGVDS 397
Db 301 QGVITVFNQVTWLSAKQHTLSDTLWPGDSRLQINFRATQPLNGRVITASFPAGVDS 360
Qy 398 PRAEPVQVNSCLAAGD 414
Db 361 PRAEPVQVNSCLAAGD 377

RESULT 14

US-09-177-776-87
Sequence 87, Application US/09177776A
Patent No. 6297051
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
FILE REFERENCE: D-0021.5A
CURRENT APPLICATION NUMBER: US/09/177,776A
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 08/787,739
EARLIER FILING DATE: 1997-01-24
EARLIER APPLICATION NUMBER: 08/485,049
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/486,756
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/477,504
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/481,658
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,862
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,863
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/487,077
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/260,190
EARLIER FILING DATE: 1994-06-15
EARLIER APPLICATION NUMBER: 08/177,093
EARLIER FILING DATE: 1993-12-30
EARLIER APPLICATION NUMBER: 07/964,589
EARLIER FILING DATE: 1992-10-21
EARLIER APPLICATION NUMBER: PV-709-92
EARLIER FILING DATE: 1992-03-11
NUMBER OF SEQ ID NOS: 116
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 87
LENGTH: 377
TYPE: PRT
ORGANISM: HUMAN
US-09-177-776-87

Query Match 82.8%; Score 2007; DB 3; Length 377;
Best Local Similarity 100.0%; Pred. No. 1,2e-161;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 38 QRLPRMQEDSPGLGGSSGDEDDPLGEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGE 97

Db 1 QRLPRMQEDSPGLGGSSGDEDDPLGEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGE 60
Qy 98 VKKSEEGSLKLEDPVTEARQDPEQNNAHKXGDDQSHNRVGGDPPMPVSPACA 157
Db 61 VKKSEEGSLKLEDPVTEARQDPEQNNAHKXGDDQSHNRVGGDPPMPVSPACA 120
Qy 158 GRQSPVDIRPOLAACCPALRPELIGFQLPELRIRNNGHSVQLTPPGLEMALGPG 217
Db 121 GRQSPVDIRPOLAACCPALRPELIGFQLPELRIRNNGHSVQLTPPGLEMALGPG 180
Qy 218 REYRALQHLHMGAGRGSEHTVEGRFPFAIHVHLSTAFARVDALGRPGGLAVLAA 277
Db 181 REYRALQHLHMGAGRGSEHTVEGRFPFAIHVHLSTAFARVDALGRPGGLAVLAA 240
Qy 278 FLEEGPENSAYEQLSLRLEIAEGSETQVPGDLISALPDSFRYQYEGSLTTPCA 337
Db 241 FLEEGPENSAYEQLSLRLEIAEGSETQVPGDLISALPDSFRYQYEGSLTTPCA 300
Qy 338 QGVITVFNQVTWLSAKQHTLSDTLWPGDSRLQINFRATQPLNGRVITASFPAGVDS 397
Db 301 QGVITVFNQVTWLSAKQHTLSDTLWPGDSRLQINFRATQPLNGRVITASFPAGVDS 360
Qy 398 PRAEPVQVNSCLAAGD 414
Db 361 PRAEPVQVNSCLAAGD 377

RESULT 15

US-07-964-589-2
Sequence 2, Application US/07964589
Patent No. 5387676
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Launder
STREET: Steuart Street Tower, 18th Fl., One Market
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/964,589
FILING DATE: 19921021
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Launder, Leona L
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-777-9257
TELEFAX: 415-543-4219
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 429 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-964-589-2

Query Match 57.7%; Score 1398.5; DB 1; Length 429;
Best Local Similarity 69.1%; Pred. No. 4e-110;
Matches 300; Conservative 12; Mismatches 83; Indels 39; Gaps 7;

```
QY 38 QRLPRMOZSDPIGGSSGEDDPLGBEDLPSEEDSPREEDPPGEEDLPGEEDLP--P 96
Db 1 QRLPRMOZSDPIGGSSGEDDPLGBEDLPSEEDSPREEDPPGEEDLPGEEDLP--P 96
QY 97 EYKPKSEEGSILKLEDLPTVEAPGDPONNARDEKGDQSHRYGSDPWP----- 150
Db 60 KLMFKSEEGSILKLEDLPTVEAPGDPONNARDEKGDQSHRYGSDPWP----- 119
QY 151 --RVSPPACAGRFQSPVDIRPOLAFC-----PALRPLELGFQPLPELRLNNGH 200
Db 120 RAASSPRWISAPSSPPSAP-----CAPMNSMASSRRSONCACQMP----- 162
QY 201 SVQTLTPGLEMALPGRE--YRALQHLHMGAGRPGSEHTVEGHRFPALIHVHLSTA 258
Db 163 --QCATDPASMARDSGRAGVPAQLHLHMGAGRPGSEHTVEGHRFPALIHVHLSTA 220
QY 259 FARVDEALGRPGGLAVLAFLBEGHEENSAYEQLLSRLEIAEGSFTQVPGDISALLP 318
Db 221 FARVDEALGRPGGLAVLAFLBEGHEENSAYEQLLSRLEIAEGSFTQVPGDISALLP 280
QY 319 SPSRYFOYEGSLTTPCAQGVITVENQTVMLSAKQHLTSLDTLWPGDSRLQNLFRAT 378
Db 281 SPSRYFOYEGSLTTPCAQGVITVENQTVMLSAKQHLTSLDTLWPGDSRLQNLFRAT 340
QY 379 QPLNGRVIBASFPAGVDSSPRAAEFVQNSCLAAGDILALVFGLLFAVTSVAFVQMRQ 438
Db 341 QPLNGRVIBASFPAGVDSSPRAAEFVQNSCLAAGDILALVFGLLFAVTSVAFVQMRQ 400
QY 439 HRRGTGGSVYRPA 452
Db 401 HRRGTGGSVYRPS 414
```

Search completed: March 5, 2004, 09:08:29
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      March 5, 2004, 09:06:36 ; Search time 40 Seconds
              (without alignments)
              2422.984 Million cell updates/sec
```

Title:	US-09-967-237-2
Perfect score:	2424
Sequence:	1 MAPLCPFWLPIIPAPAG.....RRGTGGVSYRPAEVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA.*

1:	/cgn2_6/ptodataa/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodataa/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodataa/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodataa/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodataa/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodataa/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodataa/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodataa/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodataa/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodataa/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodataa/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodataa/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodataa/1/pubpaa/US10B_PUBCOMB.pep.*
14:	/cgn2_6/ptodataa/1/pubpaa/US10S_PUBCOMB.pep.*
15:	/cgn2_6/ptodataa/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodataa/1/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodataa/1/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodataa/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2424	100.0	459	9	US-09-772-719-2	Sequence 2, Appl1
2	2424	100.0	458	10	US-09-967-237-2	Sequence 2, Appl1
3	2424	100.0	458	14	US-10-301-822-12	Sequence 12, Appl1
4	2424	100.0	459	15	US-10-463-572-10	Sequence 10, Appl
5	2424	100.0	459	15	US-10-295-027-306	Sequence 106, Appl
6	2424	100.0	458	15	US-10-295-027-1239	Sequence 1239, Appl
7	2419	99.8	610	9	US-09-783-708-1	Sequence 1, Appl1
8	2007	82.8	377	10	US-09-967-237-87	Sequence 87, Appl
9	1370	56.5	257	10	US-09-967-237-51	Sequence 51, Appl
10	1364	56.3	256	9	US-09-772-719-51	Sequence 51, Appl
11	884	36.5	170	9	US-09-772-719-54	Sequence 54, Appl
12	562	23.2	337	10	US-09-946-374-423	Sequence 423, Appl
13	562	23.2	337	10	US-09-983-000A-27	Sequence 27, Appl
14	562	23.2	337	13	US-10-006-867-74	Sequence 74, Appl
15	562	23.2	337	13	US-10-002-566-268	Sequence 268, Appl

16	562	23.2	337	13	US-10-063-594-74	Sequence 74, Appl
17	562	23.2	337	14	US-10-174-559-268	Sequence 268, Appl
18	562	23.2	337	14	US-10-176-758-268	Sequence 268, Appl
19	562	23.2	337	14	US-10-175-73-268	Sequence 268, Appl
20	562	23.2	337	14	US-10-063-616-74	Sequence 74, Appl
21	562	23.2	337	14	US-10-173-706-268	Sequence 268, Appl
22	562	23.2	337	14	US-10-175-738-268	Sequence 268, Appl
23	562	23.2	337	14	US-10-175-755-268	Sequence 268, Appl
24	562	23.2	337	14	US-10-176-485-268	Sequence 268, Appl
25	562	23.2	337	14	US-10-176-757-268	Sequence 268, Appl
26	562	23.2	337	14	US-10-176-912-268	Sequence 268, Appl
27	562	23.2	337	14	US-10-180-555-268	Sequence 268, Appl
28	562	23.2	337	14	US-10-180-555-268	Sequence 268, Appl
29	562	23.2	337	14	US-10-063-502-74	Sequence 74, Appl
30	562	23.2	337	14	US-10-173-700-268	Sequence 268, Appl
31	562	23.2	337	14	US-10-174-572-268	Sequence 268, Appl
32	562	23.2	337	14	US-10-174-579-268	Sequence 268, Appl
33	562	23.2	337	14	US-10-174-589-268	Sequence 268, Appl
34	562	23.2	337	14	US-10-174-588-268	Sequence 268, Appl
35	562	23.2	337	14	US-10-175-739-268	Sequence 268, Appl
36	562	23.2	337	14	US-10-175-744-268	Sequence 268, Appl
37	562	23.2	337	14	US-10-175-742-268	Sequence 268, Appl
38	562	23.2	337	14	US-10-176-488-268	Sequence 268, Appl
39	562	23.2	337	14	US-10-176-499-268	Sequence 268, Appl
40	562	23.2	337	14	US-10-176-747-268	Sequence 268, Appl
41	562	23.2	337	14	US-10-176-750-268	Sequence 268, Appl
42	562	23.2	337	14	US-10-176-985-268	Sequence 268, Appl
43	562	23.2	337	14	US-10-176-987-268	Sequence 268, Appl
44	562	23.2	337	14	US-10-176-992-268	Sequence 268, Appl
45	562	23.2	337	14	US-10-176-993-268	Sequence 268, Appl

ALIGNMENTS

RESULT 1
; Sequence 2, Application US/09772719
; Patent No. US20020137910A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/772,719
; FILING DATE: 30-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
;

Db 1 MAPLCPSPWLLPILIPAPAGLTVQLLLSLLLMFVHQRLLPRMGEDSLGGSSGGBDPL 60
Qy 61 GEEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Qy 121 DPOEPONNARXKEDDQSHMRVYGGDPMPRVSPACAGRFOSPDYIRPOLAACPAPRL 180
Db 121 DPOEPONNARXKEDDQSHMRVYGGDPMPRVSPACAGRFOSPDYIRPOLAACPAPRL 180
Qy 181 ELIGFOLPPLPELRLRNNGHSVQTLPLPGLMAGPGRREYALQLHLHWGAAGRPGSEHT 240
Db 181 ELIGFOLPPLPELRLRNNGHSVQTLPLPGLMAGPGRREYALQLHLHWGAAGRPGSEHT 240
Qy 241 VEGHRRPAEIHVHLSSTARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLEIEA 300
Db 241 VEGHRRPAEIHVHLSSTARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLEIEA 300
Qy 301 EEGSETQVPGDLISALLPSDFSRFYQEGSLTTPPCAQGIWTVFNQTVWLSAKQHTLS 360
Db 301 EEGSETQVPGDLISALLPSDFSRFYQEGSLTTPPCAQGIWTVFNQTVWLSAKQHTLS 360
Qy 361 DTLMGPGDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAEPVOLNSCLAAGDILALVF 420
Db 361 DTLMGPGDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAEPVOLNSCLAAGDILALVF 420
Qy 421 GLLFAVTSVAFVQMRROHRRGTGGSYRPAEVAETGA 459
Db 421 GLLFAVTSVAFVQMRROHRRGTGGSYRPAEVAETGA 459

RESULT 4
US-10-465-572-10
; Sequence 10, Application US/10465572
; Publication No. US20030207840A1
; GENERAL INFORMATION:
; APPLICANT: Rigging, Gregory
; APPLICANT: Lal, Anita
; TITLE OF INVENTION: GENES INDUCED BY HYPOXIA
; FILE REFERENCE: 000250.00012
; CURRENT APPLICATION NUMBER: US/10/465.572
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US/10/201.642
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/307.600
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-465-572-10

Query Match 100.0%; Score 2424; DB 15; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.6e-181;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPLCPSPWLLPILIPAPAGLTVQLLLSLLLMFVHQRLLPRMGEDSLGGSSGGBDPL 60
Db 1 MAPLCPSPWLLPILIPAPAGLTVQLLLSLLLMFVHQRLLPRMGEDSLGGSSGGBDPL 60
Qy 61 GEEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Qy 121 DPOEPONNARXKEDDQSHMRVYGGDPMPRVSPACAGRFOSPDYIRPOLAACPAPRL 180
Db 121 DPOEPONNARXKEDDQSHMRVYGGDPMPRVSPACAGRFOSPDYIRPOLAACPAPRL 180
Qy 181 ELIGFOLPPLPELRLRNNGHSVQTLPLPGLMAGPGRREYALQLHLHWGAAGRPGSEHT 240
Db 181 ELIGFOLPPLPELRLRNNGHSVQTLPLPGLMAGPGRREYALQLHLHWGAAGRPGSEHT 240

Db 181 ELIGFOLPPLPELRLRNNGHSVQTLPLPGLMAGPGRREYALQLHLHWGAAGRPGSEHT 240
Qy 241 VEGHRRPAEIHVHLSSTARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLEIEA 300
Db 241 VEGHRRPAEIHVHLSSTARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLEIEA 300
Qy 301 EEGSETQVPGDLISALLPSDFSRFYQEGSLTTPPCAQGIWTVFNQTVWLSAKQHTLS 360
Db 301 EEGSETQVPGDLISALLPSDFSRFYQEGSLTTPPCAQGIWTVFNQTVWLSAKQHTLS 360
Qy 361 DTLMGPGDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAEPVOLNSCLAAGDILALVF 420
Db 361 DTLMGPGDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAEPVOLNSCLAAGDILALVF 420
Qy 421 GLLFAVTSVAFVQMRROHRRGTGGSYRPAEVAETGA 459
Db 421 GLLFAVTSVAFVQMRROHRRGTGGSYRPAEVAETGA 459

RESULT 5
US-10-295-027-306
; Sequence 306, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hewez, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295.027
; CURRENT FILING DATE: 2002-11-13/733
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 306
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-306

Query Match 100.0%; Score 2424; DB 15; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.6e-181;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPLCPSPWLLPILIPAPAGLTVQLLLSLLLMFVHQRLLPRMGEDSLGGSSGGBDPL 60

```

Db      1 MAPLCSPMWPLPLIAPAPGLTYQLLSLLMLMVFHQRLPRMQEDSLPGGSSGSDPDL 60
Qy      61 GREDLPSEEDSPREDDPGSEDLPGSEDLPGSEDLPEVKKSEEGSLKLEDLPTVEARG 120
Db      61 GREDLPSEEDSPREDDPGSEDLPGSEDLPGSEDLPEVKKSEEGSLKLEDLPTVEARG 120
Qy      121 DPOEPQNNNAHRKEDDQSHMRVGGDPMPMPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Db      121 DPOEPQNNNAHRKEDDQSHMRVGGDPMPMPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Qy      181 ELIGFOLPPLPELRLRNNGHSVOLTLPPELEMAIGREYRALQHLHMGAAGRPSEHT 240
Db      181 ELIGFOLPPLPELRLRNNGHSVOLTLPPELEMAIGREYRALQHLHMGAAGRPSEHT 240
Qy      241 VEGHRFPAPAEIHVYHLSTAFARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLEEA 300
Db      241 VEGHRFPAPAEIHVYHLSTAFARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLEEA 300
Qy      301 EGGSETQVPGDLISALPSPDSRYFOYEGSLTTPPCAQGVIMVFQOTWMLSAKOLHTIS 360
Db      301 EGGSETQVPGDLISALPSPDSRYFOYEGSLTTPPCAQGVIMVFQOTWMLSAKOLHTIS 360
Qy      361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPAAEPVQNSCLAAGDILALVF 420
Db      361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPAAEPVQNSCLAAGDILALVF 420
Qy      421 GLFPAVTSVAFVQMRQHRGTGKGVSYRPAEVAETGA 459
Db      421 GLFPAVTSVAFVQMRQHRGTGKGVSYRPAEVAETGA 459

RESULT 6
US-10-295-027-1239
; Sequence 1239, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natshea
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.

```

```

; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1239
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1239
Query Match      100.0%; Score 2424; DB 15; Length 459;
Best Local Similarity 100.0%; Pred. No. 1,6e-181;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MAPLCSPMWPLPLIAPAPGLTYQLLSLLMLMVFHQRLPRMQEDSLPGGSSGSDPDL 60
Db      1 MAPLCSPMWPLPLIAPAPGLTYQLLSLLMLMVFHQRLPRMQEDSLPGGSSGSDPDL 60
Qy      61 GREDLPSEEDSPREDDPGSEDLPGSEDLPGSEDLPEVKKSEEGSLKLEDLPTVEARG 120
Db      61 GREDLPSEEDSPREDDPGSEDLPGSEDLPGSEDLPEVKKSEEGSLKLEDLPTVEARG 120
Qy      121 DPOEPQNNNAHRKEDDQSHMRVGGDPMPMPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Db      121 DPOEPQNNNAHRKEDDQSHMRVGGDPMPMPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Qy      181 ELIGFOLPPLPELRLRNNGHSVOLTLPPELEMAIGREYRALQHLHMGAAGRPSEHT 240
Db      181 ELIGFOLPPLPELRLRNNGHSVOLTLPPELEMAIGREYRALQHLHMGAAGRPSEHT 240
Qy      241 VEGHRFPAPAEIHVYHLSTAFARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLEEA 300
Db      241 VEGHRFPAPAEIHVYHLSTAFARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLEEA 300
Qy      301 EGGSETQVPGDLISALPSPDSRYFOYEGSLTTPPCAQGVIMVFQOTWMLSAKOLHTIS 360
Db      301 EGGSETQVPGDLISALPSPDSRYFOYEGSLTTPPCAQGVIMVFQOTWMLSAKOLHTIS 360
Qy      361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPAAEPVQNSCLAAGDILALVF 420
Db      361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPAAEPVQNSCLAAGDILALVF 420
Qy      421 GLFPAVTSVAFVQMRQHRGTGKGVSYRPAEVAETGA 459
Db      421 GLFPAVTSVAFVQMRQHRGTGKGVSYRPAEVAETGA 459

```

```

RESULT 7
US-09-783-708-1
; Sequence 1, Application US/09783708
; Patent No. US20020058041A1
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: KIDNEY-SPECIFIC TUMOR VACCINE DIRECTED AGAINST KIDNEY TUMOR ANT
; FILE REFERENCE: 3067-897420US
; CURRENT APPLICATION NUMBER: US/09/783,708
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/182,429
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/182,636
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: G250-GM-CSF fusion protein
US-09-783-708-1

```

```

Query Match      99.8%; Score 2419; DB 9; Length 610;
Best Local Similarity 100.0%; Pred. No. 5,6e-181;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 2 APLCPSPWLPPLIPAPAGLTVQLLILLIMVHPORLPRMOEDSPLOGGSGGDDPLG 61
DB 147 APLCPSPWLPPLITAPAPAGLTVQLLILLIMVHPORLPRMOEDSPLOGGSGGDDPLG 206
QY 62 EEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 121
DB 207 EEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 266
QY 122 POEPONNAHRKEDDQSHWRVYSGDPMPVSPACAGFQSPVDIRPOLAFCPALRPE 181
DB 267 POEPONNAHRKEDDQSHWRVYSGDPMPVSPACAGFQSPVDIRPOLAFCPALRPE 326
QY 182 LLGFQPLPLPELRKLNNGHSVQLTLPPGLEMALGPGREYRALQHLHMGAGRPGSEHTV 241
DB 327 LLGFQPLPLPELRKLNNGHSVQLTLPPGLEMALGPGREYRALQHLHMGAGRPGSEHTV 386
QY 242 ECHRPFAITHVHSTAFARVDEALGRPGGLAVLAFLLEGPEENSAYEQLLSRLBEIAE 301
DB 387 ECHRPFAITHVHSTAFARVDEALGRPGGLAVLAFLLEGPEENSAYEQLLSRLBEIAE 446
QY 302 EGSETQVPGDLISALLPSDFSRVYQYEGSLTTPPCAQGVIMVFNQTMLSAKQHTLSD 361
DB 447 EGSETQVPGDLISALLPSDFSRVYQYEGSLTTPPCAQGVIMVFNQTMLSAKQHTLSD 506
QY 362 TLMGDSRLQUNFRATQPLNGRVIEASFPAGVDSPPRAEPVQUNSCLAAGDILALVFG 421
DB 507 TLMGDSRLQUNFRATQPLNGRVIEASFPAGVDSPPRAEPVQUNSCLAAGDILALVFG 566
QY 422 LLFAVTSVAFVOMRQHRGTGKGVSYRPAEVAETGA 459
DB 567 LLFAVTSVAFVOMRQHRGTGKGVSYRPAEVAETGA 604
```

RESULT 8

```
US-09-967-237-87
; Sequence 87, Application US/09967237
; Publication No. US20030049828A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5B-2
; CURRENT APPLICATION NUMBER: US/09/967,237
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 09/178,115
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 87
; LENGTH: 377
; TYPE: PRT
; ORGANISM: HUMAN
US-09-967-237-87
```

```
Query Match 82.8%; Score 2007; DB 10; Length 377;
Best Local Similarity 100.0%; Pred. No. 5.6e-149;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 38 ORLPKMOEDSPLOGGSGGDDPLGEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLP 97
DB 1 ORLPKMOEDSPLOGGSGGDDPLGEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLP 60
QY 98 VKPKSEEGSLKLEDLPTVEAPGDPOEPONNAHRKEDDQSHWRVYSGDPMPVSPACA 157
DB 61 VKPKSEEGSLKLEDLPTVEAPGDPOEPONNAHRKEDDQSHWRVYSGDPMPVSPACA 120
QY 158 GFGQSPVDIRPOLAFCPALRPELTLGGFQPLPELRKLNNGHSVQLTLPPGLEMALGPG 217
DB 121 GFGQSPVDIRPOLAFCPALRPELTLGGFQPLPELRKLNNGHSVQLTLPPGLEMALGPG 180
QY 218 REYRALQHLHMGAGRPGSEHTVEGHRPFAEIHVHLSFAFVDEALGRPGGLAVLAA 277
```

```
DB 181 REYRALQHLHMGAGRPGSEHTVEGHRPFAEIHVHLSFAFVDEALGRPGGLAVLAA 240
QY 278 FLEGPEENSAYEQQLLSRLBEIAEGSETOVPGDLISALLPSDFSRVYQYEGSLTTPPCA 337
DB 241 FLEGPEENSAYEQQLLSRLBEIAEGSETOVPGDLISALLPSDFSRVYQYEGSLTTPPCA 300
QY 338 QGVIMVFNQTMLSAKQHTLSDITLMGDSRLQUNFRATQPLNGRVIEASFPAGVDS 397
DB 301 QGVIMVFNQTMLSAKQHTLSDITLMGDSRLQUNFRATQPLNGRVIEASFPAGVDS 360
QY 398 PRAEPVQUNSCLAAGD 414
DB 361 PRAEPVQUNSCLAAGD 377
```

RESULT 9

```
US-09-967-237-51
; Sequence 51, Application US/09967237
; Publication No. US20030049828A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5B-2
; CURRENT APPLICATION NUMBER: US/09/967,237
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 09/178,115
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 51
; LENGTH: 257
; TYPE: PRT
; ORGANISM: HUMAN
US-09-967-237-51
```

```
Query Match 56.5%; Score 1370; DB 10; Length 257;
Best Local Similarity 100.0%; Pred. No. 2.7e-99;
Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 135 GDDDSHWRVYSGDPMPVSPACAGRFQSPVDIRPOLAFCPALRPELTLGGFQPLPELR 194
DB 1 GDDDSHWRVYSGDPMPVSPACAGRFQSPVDIRPOLAFCPALRPELTLGGFQPLPELR 60
QY 195 LRNNGHSVQLTLPPGLEMALGPGREYRALQHLHMGAGRPGSEHTVEGHRPFAEIHVH 254
DB 61 LRNNGHSVQLTLPPGLEMALGPGREYRALQHLHMGAGRPGSEHTVEGHRPFAEIHVH 120
QY 255 LSTAFARVDEALGRPGGLAVLAFLLEGPEENSAYEQQLLSRLBEIAEGSETOVPGDLIS 314
DB 121 LSTAFARVDEALGRPGGLAVLAFLLEGPEENSAYEQQLLSRLBEIAEGSETOVPGDLIS 180
QY 315 ALLPSDFSRVYQYEGSLTTPPCAQGVIMVFNQTMLSAKQHTLSDITLMGDSRLQUN 374
DB 181 ALLPSDFSRVYQYEGSLTTPPCAQGVIMVFNQTMLSAKQHTLSDITLMGDSRLQUN 240
QY 375 FRATQPLNGRVIEASFP 391
DB 241 FRATQPLNGRVIEASFP 257
```

RESULT 10

```
US-09-772-719-51
; Sequence 51, Application US/09772719
; Patent No. US20020137910A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
```

```

ADDRESS: Leona L. Lauder
STREET: 369 Pine Street
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/772,719
FILING DATE: 30-JAN-2001
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: carbonic anhydrase domain
US-09-772-719-51

Query Match          56.3%; Score 1364; DB 9; Length 256;
Best Local Similarity 100.0%; Pred. No. 7.8e-99;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 DDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPPELLGFOUPPELR 195
DB 1 DDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPOLAFCPALRPPELLGFOUPPELR 60
QY 196 RNNHGSVQLTPPCLLENALGPGREYRALQLHLWGAAAGRGSGHTVEGRFPALHVVHL 255
DB 61 RNNHGSVQLTPPCLLENALGPGREYRALQLHLWGAAAGRGSGHTVEGRFPALHVVHL 120
QY 256 STAFARVDEALGRPGGLAVLAAPLEEGPEENSAYEQLLSRLEEIAEGSETOVPGIDISA 315
DB 121 STAFARVDEALGRPGGLAVLAAPLEEGPEENSAYEQLLSRLEEIAEGSETOVPGIDISA 180
QY 316 LLPDFFRFFQYEGSLTTPPCAGQVITWVFNQVWMSAKQLHTLSLTWNGPDSRLQLNF 375
DB 181 LLPDFFRFFQYEGSLTTPPCAGQVITWVFNQVWMSAKQLHTLSLTWNGPDSRLQLNF 240
QY 376 RATOPNGRVIEASFP 391
DB 241 RATOPNGRVIEASFP 256

```

```

STREET: 369 Pine Street
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/772,719
FILING DATE: 30-JAN-2001
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-772-719-54

Query Match          36.5%; Score 884; DB 9; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.8e-61;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 RALQLHLWGAAAGRPSEHTVEGRFPALHVVHLSTAFARVDEALGRPGGLAVLAAPLE 280
DB 1 RALQLHLWGAAAGRPSEHTVEGRFPALHVVHLSTAFARVDEALGRPGGLAVLAAPLE 60
QY 281 EGPENSAYEQLLSRLEEIAEGSETOVPGIDISALPDSRFFQYEGSLTTPPCAGV 340
DB 61 EGPENSAYEQLLSRLEEIAEGSETOVPGIDISALPDSRFFQYEGSLTTPPCAGV 120
QY 341 IWTVFNQVWMSAKQLHTLSDTLWNGPDSRLQLNFRATQPLNGRVIEASF 390
DB 121 IWTVFNQVWMSAKQLHTLSDTLWNGPDSRLQLNFRATQPLNGRVIEASF 170

```

```

RESULT 11
US-09-772-719-54
Sequence 54, Application US/09772719
Patent No. US20020137910A1
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSER: Leona L. Lauder

```

```

RESULT 12
US-09-966-374-423
Sequence 423, Application US/09946374
Publication No. US20030073129A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth U.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.

```

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2830P1C1
CURRENT APPLICATION NUMBER: US/09/946,374
CURRENT FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 60/098716
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098723
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098750
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098803
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098821
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098843
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/099536
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099586
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099602
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099642
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099741
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099754
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099763
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099792
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099808
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099815
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100388
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100584
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100661
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100662
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100664
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100683
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100684
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100710
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100711
PRIOR FILING DATE: 1998-09-17

PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100849
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101014
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101068
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101071
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101279
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/101471
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101472
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101474
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101475
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101476
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101479
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101743
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101915
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/102207
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102240
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102307
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102330
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102331
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102484
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102487
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102570
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102571
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102684
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/102687
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/102965
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 60/103258
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103314
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103315
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103328
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103395

PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103396
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103401
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: 60/103449
 PRIOR FILING DATE: 1998-10-06
 PRIOR APPLICATION NUMBER: 60/103633
 PRIOR FILING DATE: 1998-10-08
 PRIOR APPLICATION NUMBER: 60/103678
 PRIOR FILING DATE: 1998-10-08
 PRIOR APPLICATION NUMBER: 60/103679
 PRIOR FILING DATE: 1998-10-08
 PRIOR APPLICATION NUMBER: 60/103711
 PRIOR FILING DATE: 1998-10-08
 PRIOR APPLICATION NUMBER: 60/104257
 PRIOR FILING DATE: 1998-10-14
 PRIOR APPLICATION NUMBER: 60/104987
 PRIOR FILING DATE: 1998-10-20
 PRIOR APPLICATION NUMBER: 60/105000
 PRIOR FILING DATE: 1998-10-20
 PRIOR APPLICATION NUMBER: 60/105002
 PRIOR FILING DATE: 1998-10-20
 PRIOR APPLICATION NUMBER: 60/105104
 PRIOR FILING DATE: 1998-10-21
 PRIOR APPLICATION NUMBER: 60/105169
 PRIOR FILING DATE: 1998-10-22
 PRIOR APPLICATION NUMBER: 60/105266
 PRIOR FILING DATE: 1998-10-22
 PRIOR APPLICATION NUMBER: 60/105693
 PRIOR FILING DATE: 1998-10-26
 PRIOR APPLICATION NUMBER: 60/105694
 PRIOR FILING DATE: 1998-10-26
 PRIOR APPLICATION NUMBER: 60/105807

Query Match 23.2%; Score 562; DB 10; Length 337;
 Best Local Similarity 38.8%; Pred. No. 7,7e-36;
 Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;

QY 136 DDQSHWRY---GGDPMPRVSPACAGRFQSPVDIRPOLAACFPALPFLLEGFOLPPLPE 192
 DB 17 DGGQHWTEGPHGQDHPAGYPCGNNAGSPIDIQDVSVPDPLPALQPHGDOQTEP 76
 QY 193 LRLRNNGHSVOLTLPGLMALGPGREYALQLHLMGAAGRP-GSEHTVEGHRFPALH 251
 DB 77 LDHNNHGTIVQLSPSTLYIG-GLPRKYVAAQHLHMGQKSGSGSEHQINSEATFAELH 135
 QY 252 VVHL-STAFARVDEALGPRGGLAVLAFLIEGSEENSAVEQLLSRLIEEISSEETQVFG 310
 DB 136 IVHYDSDSYDSLSEAEERPOGLAVLGILIEVGETKNIAVEHLISLHEVRHKDKTSVPP 195
 QY 311 LDISALLPDSFSGRYPOEGSLTPPCAGQVITVFNQVTWLSAKQLHTLSDTLNGPDSR 370
 DB 196 FNRLBELPKQGLQGYFRYNSLTPPCYQSVLMTVYFRSGQISMEQLHQLQTLSTEEPP 255
 QY 371 LQI---NFRATQPLNGRVIIEASFPAGVSSPRAAEPVQNSCLAAGDILALVFGLLFA-- 425
 DB 256 SKLVQNVYRALQPLNQHWVFASF-----IQAGSYTTGEMLSLGVGILVGL 302
 QY 426 --VTSYALVQMRQRHRTGKGVSYRAE 453
 DB 303 CLLAVYFTARKIRKRLKRLKRSVVTSAQ 332

RESULT 13
 US-09-963-000A-27
 Sequence 27, Application US/09963000A
 Publication No. US20030118585A1

GENERAL INFORMATION:
 APPLICANT: AGY Therapeutics
 APPLICANT: Melcher, Thorsten
 APPLICANT: Mueller, Sabine
 APPLICANT: Chin, Daniel

TITLE OF INVENTION: USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VIST
 TITLE OF INVENTION: OF BRAIN TUMORS
 FILE REFERENCE: 263/180 -- Peagleman -- AGY
 CURRENT APPLICATION NUMBER: US/09/963,000A
 CURRENT FILING DATE: 2001-10-17
 NUMBER OF SEQ ID NOS: 35
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 27
 LENGTH: 337
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: Gene
 LOCATION: (1)..(337)
 OTHER INFORMATION: Carbonic anhydrase domain of human carbonic anhydrase VIX
 US-09-963-000A-27

Query Match 23.2%; Score 562; DB 10; Length 337;
 Best Local Similarity 38.8%; Pred. No. 7,7e-36;
 Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;

QY 136 DDQSHWRY---GGDPMPRVSPACAGRFQSPVDIRPOLAACFPALPFLLEGFOLPPLPE 192
 DB 17 DGGQHWTEGPHGQDHPAGYPCGNNAGSPIDIQDVSVPDPLPALQPHGDOQTEP 76
 QY 193 LRLRNNGHSVOLTLPGLMALGPGREYALQLHLMGAAGRP-GSEHTVEGHRFPALH 251
 DB 77 LDHNNHGTIVQLSPSTLYIG-GLPRKYVAAQHLHMGQKSGSGSEHQINSEATFAELH 135
 QY 252 VVHL-STAFARVDEALGPRGGLAVLAFLIEGSEENSAVEQLLSRLIEEISSEETQVFG 310
 DB 136 IVHYDSDSYDSLSEAEERPOGLAVLGILIEVGETKNIAVEHLISLHEVRHKDKTSVPP 195
 QY 311 LDISALLPDSFSGRYPOEGSLTPPCAGQVITVFNQVTWLSAKQLHTLSDTLNGPDSR 370
 DB 196 FNRLBELPKQGLQGYFRYNSLTPPCYQSVLMTVYFRSGQISMEQLHQLQTLSTEEPP 255
 QY 371 LQI---NFRATQPLNGRVIIEASFPAGVSSPRAAEPVQNSCLAAGDILALVFGLLFA-- 425
 DB 256 SKLVQNVYRALQPLNQHWVFASF-----IQAGSYTTGEMLSLGVGILVGL 302
 QY 426 --VTSYALVQMRQRHRTGKGVSYRAE 453
 DB 303 CLLAVYFTARKIRKRLKRLKRSVVTSAQ 332

RESULT 14
 US-10-006-867-74
 Sequence 74, Application US/10006867
 Publication No. US20020119130A1

GENERAL INFORMATION:
 APPLICANT: Baton, Dan L.
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, Christopher J.
 APPLICANT: Gurney, Auecin L.
 APPLICANT: Watanabe, Colin K.
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3230R1C1
 CURRENT APPLICATION NUMBER: US/10/006,867
 CURRENT FILING DATE: 2001-12-06
 PRIOR APPLICATION NUMBER: 60/063435
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/064215
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/082797
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/083495
 PRIOR FILING DATE: 1998-04-29

[illegible]

PRIOR APPLICATION NUMBER: 60/120014
PRIOR FILING DATE: 1999-02-11
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/129674
PRIOR FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/138387
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/163495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/175481
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 60/191007
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/199397
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380142

Query Match

23.2%; Score 562; DB 13; Length 337;

Best Local Similarity 38.8%; Pred. No. 7,7e-36;
Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;

QY 136 DDGSHRY--GDBPWRVSPACAGRFQSPVDIRPOLAFCAALRELLGQPLPE 192
DB 17 DGGQHTYEGPHGQDHPWASYFBCGNNASPIDQDTSVTFDDPLPQPHGYDQCTEP 76
QY 193 LRLRNNGHVSQVLTTPPGLEMALGPGREYRALQHLHNGAARP-GSEHTVEGHRPAEIH 251
DB 77 LDHNNHGTQLSJPSSTLYLG-GLPRKYVAAQHLHFWGQSGSGSHQNSBATPAEIH 135
QY 252 VVHL-STAFARVDLHGPGGLAVLAFLEREPENAVYQQLSRLEIEBSEFQVQG 310
DB 136 IVHYDSYDLSSEAERPQGLAVLGLIEVGETKNLAYEHILSHLEVHKQKTSVFP 195
QY 311 LDISALPDSFSRYFQYEGSLTTPPCAQGVITWFENQTVMLSAKQHTLSDTLGPGDSR 370
DB 196 FHLRELPLKQJGQYRVNGSLTTPCYQSVLMTVFYRSGQISMEQLEKQGLFSTEEBP 255
QY 371 LQI--NFRATQPLNGRVTEASFPAGVDSSFRAAEPVQUNSCLAAGDILALVFGCLFA-- 425
DB 256 SKLLVQNRALQPLQKRVFASF-----IQAGSYTTGEMLSLGVGILVGC 302
QY 426 --VTSVAFVQWRQHRGRTGSGVSRPAE 453
DB 303 CILLAVYFIARKIRKRIENRKSXYVTSAQ 332

RESULT 15

US-10-052-586-268

Sequence 268, Application US/10052586

Publication No. US20020127584A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/074450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/081070
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081195
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081838
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082568
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082569
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082704
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082797
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083495
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083496
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083499
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/086023
 PRIOR FILING DATE: 1998-05-18
 PRIOR APPLICATION NUMBER: 60/086392
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/086486
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087098
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087208
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212

PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088722
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088740
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088811
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088825
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088863
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089090
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089908

Query Match 23.2%; Score 562; DB 13; Length 337;
 Best Local Similarity 38.8%; Pred. No. 7.7e-36;
 Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;

QY 136 DQSHWRY---GDBPWVRVSPACAGRFQSPVDIRPOLAFCPALRPLELGFOLPLPE 192
 DB 17 DGGQWTEGPHGQDHWPAASYPCGNNASQSTIDQDTSVTFDPDLPALOPHGIDQPTET 76
 QY 193 LRLRNNGHVSQVLTTPGJEMALGPGREYRALQHLHWGAARP-GSEHTVEGHRFPATEH 251
 DB 77 LDHNNHGHVQLSLPSTLYLG-GLPRKYVAAQLHLHWGQKSGPSGSEHQINSBATFAELH 135
 QY 252 VYHL-STPAKYDEALGPGGLATVLAFLBEPBENAYQQLSRLEIABESSETPYRG 310
 DB 136 IYHYDSIDYDLSSEAERPGGLAVGLILEVGETKNLAYHITLSHLEVHKQKQSVPP 195
 QY 311 LDISALPSDFSRVQYEGSLTTPCAQGVMTVFNQVWLSAKQTLTSLDTLWPGDSR 370
 DB 196 FVLRRLBLKQQLGYRYNGSLTTPCYQSVMTVFRYSQISWQLEKLGTLFSTEEP 255
 QY 371 LQI---NFRATQPLNGRYTEASFPAGYDSSPRAAPVQUNSCLAAGDILVGLFLA-- 425
 DB 256 SKLIVQNYRALQPLNQRFVASF-----IQAGSSYTTGMLSLGVGILWGL 302
 QY 426 --VTSVAFVQWRQHRGRGTGQGVSYRPAE 453
 DB 303 CULAVYFIAKIRIKRKLNNKSVFTISAQ 332

Mon Mar 8 10:16:31 2004

us-09-967-237-2.rapb

Page 12

Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2004, 09:00:14 ; Search time 18 Seconds

(without alignments)
1327.789 Million cell updates/sec

Title: US-09-967-237-2

Perfect score: 2424
Sequence: 1 MAPLCPSPMLPLIPAPAPG.....RRGTGVSYPRAVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	459	CAH9_HUMAN	Q16790 homo sapien
2	1596	65.8	437	CAH9_MOUSE	Q8VBD5 mus musculu
3	562	23.2	337	CAH9_HUMAN	Q9UUX7 homo sapien
4	558	23.0	337	CAH9_MOUSE	Q9UUX7 mus musculu
5	537.5	22.2	354	CAH9_HUMAN	Q43570 homo sapien
6	526	21.7	307	CAH9_MOUSE	Q08600 ovls aries
7	512	21.1	355	CAH9_MOUSE	Q08600 ovls aries
8	511	21.1	319	CAH9_MOUSE	Q08600 ovls aries
9	496	20.5	354	CAH9_MOUSE	Q08600 ovls aries
10	488	20.1	308	CAH9_MOUSE	Q08600 ovls aries
11	467.5	19.3	317	CAH9_MOUSE	Q08600 ovls aries
12	422.5	17.4	324	CAH9_MOUSE	Q08600 ovls aries
13	422	17.4	259	CAH9_MOUSE	Q08600 ovls aries
14	419.5	17.3	290	CAH9_MOUSE	Q08600 ovls aries
15	412	17.0	289	CAH9_MOUSE	Q08600 ovls aries
16	409	16.9	264	CAH9_MOUSE	Q08600 ovls aries
17	402.5	16.6	260	CAH9_MOUSE	Q08600 ovls aries
18	398.5	16.4	255	CAH9_MOUSE	Q08600 ovls aries
19	391	16.1	259	CAH9_MOUSE	Q08600 ovls aries
20	379.5	15.7	259	CAH9_MOUSE	Q08600 ovls aries
21	379.5	15.7	259	CAH9_MOUSE	Q08600 ovls aries
22	379	15.6	259	CAH9_MOUSE	Q08600 ovls aries
23	376	15.6	262	CAH9_MOUSE	Q08600 ovls aries
24	376	15.5	262	CAH9_MOUSE	Q08600 ovls aries
25	370	15.3	260	CAH9_MOUSE	Q08600 ovls aries
26	368.5	15.2	259	CAH9_MOUSE	Q08600 ovls aries
27	368.5	15.2	259	CAH9_MOUSE	Q08600 ovls aries
28	366.5	15.1	260	CAH9_MOUSE	Q08600 ovls aries
29	364.5	15.0	259	CAH9_MOUSE	Q08600 ovls aries
30	362.5	15.0	260	CAH9_MOUSE	Q08600 ovls aries
31	359.5	14.8	260	CAH9_MOUSE	Q08600 ovls aries
32	354	14.6	260	CAH9_MOUSE	Q08600 ovls aries
33	353.5	14.6	259	CAH9_MOUSE	Q08600 ovls aries

34	352	14.5	262	1	CAH1_MONDO	Q8HY33 monodelphis
35	352	14.5	305	1	CAH4_MOUSE	Q64444 mus musculu
36	350.5	14.5	259	1	CAH3_MOUSE	P16015 mus musculu
37	347	14.3	309	1	CAH4_RAT	P48284 rattus norv
38	342.5	14.1	259	1	CAH3_HUMAN	P07451 homo sapien
39	342	14.1	304	1	CAH5_RAT	P43165 rattus norv
40	341	14.1	260	1	CAH1_SHEEP	P48282 ovls aries
41	340.5	14.0	317	1	CAH5_HUMAN	Q9Y260 homo sapien
42	339.5	14.0	317	1	CAH5_MOUSE	Q9G260 mus musculu
43	338.5	14.0	259	1	CAH3_HORSE	P07450 equus cabal
44	334	13.8	309	1	CAH5_CAEEL	Q10462 caenorhabdi
45	333	13.7	308	1	CAH4_RABIT	P48283 oryctolagus

ALIGNMENTS

RESULT 1	CAH9_HUMAN	STANDARD	PRT	459 AA.
ID	CAH9_HUMAN			
AC	Q16750			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Carbonic anhydrase IX precursor (BC 4.2.1.1) (carbonate dehydratase IX) (CA-IX) (CAIX) (Membrane antigen MN) (P54/58N) (Renal cell carcinoma-associated antigen G250) (RCC-associated antigen G250) (PMN1).			
DE	CA9 OR MN OR G250.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
CC	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RC	TISSUE=Carcinoma;			
RX	MEDLINE=9436734; PubMed=8084592;			
RA	Pastorek J., Pastorekova S., Callebaut I., Mornon J.-P., Zelink V., Oavsky R., Zet, Ovicova M., Liao S., Portetelle D., Stanbridge E.J., Zavada J., Burny A., Kettmann R.,			
RA	"Cloning and characterization of MN, a human tumor-associated protein with a domain homologous to carbonic anhydrase and a putative helix-loop-helix DNA binding segment."			
RT	Oncogene 9:2877-2888(1994).			
RL	[2]			
RP	SEQUENCE FROM N.A., AND VARIANT VAL-33.			
RC	TISSUE=Renal cell carcinoma;			
RX	MEDLINE=20175484; PubMed=10709109;			
RA	Grabmaier K., Vissers J.L.M., De Weijert M.C.A., Oosterwijk-Wakka J.C., Van Bokhoven A., Brakenhoff R.H., Noessner E., Mulders P.A., Merks G., Figdor C.G., Adema G.J., Oosterwijk E.,			
RA	"Molecular cloning and immunogenicity of renal cell carcinoma-associated antigen G250."			
RT	Int. J. Cancer 85:865-870(2000).			
RL	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Colon;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettmann M., Madan A.C., Shevchenko Y., Bouffard G.G., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,			

RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (4)
 RN CHARACTERIZATION.
 RX MEDLINE=92188536; PubMed=1312272;
 RX Pastorekova S., Zavadova Z., Kostal M., Babusikova O., Zavadova J.;
 RT "A novel quasi-viral agent, Matu, is a two-component system";
 RL Virology 187:620-626(1992).
 RN [5]
 RN SUBCELLULAR LOCATION.
 RX MEDLINE=93252505; PubMed=8486430;
 RA Zavadova J., Zavadova Z., Pastorekova S., Ciampor F., Pastorek J.,
 RA Zelnik V.;
 RT "Expression of Matu-MN protein in human tumor cultures and in clinical
 RT specimens";
 RL Int. J. Cancer 54:268-274(1993).
 CC -1- FUNCTION: Reversible hydration of carbon dioxide. May be involved
 CC in the control of cell proliferation and transformation. Appears
 CC to be a novel specific biomarker for a cervical neoplasia.
 CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
 CC -1- COFACTOR: Zinc (By similarity).
 CC -1- SUBUNIT: Forms oligomers linked by disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: FOUND ON THE SURFACE MICROVILLI AND IN THE
 CC NUCLEUS, PARTICULARLY IN NOCCELOS.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PRIMARILY IN CARCINOMA CELLS LINES.
 CC EXPRESSION IS RESTRICTED TO VERY FEW NORMAL TISSUES AND THE MOST
 CC ABUNDANT EXPRESSION IS FOUND IN THE EPITHELIAL CELLS OF GASTRIC
 CC MUCOSA.
 CC -1- PTM: N-glycosylated.
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isdb-sib.ch/announce/>
 CC or send an email to license@sib-sb.ch).
 CC -----
 DR EMBL; X66839; CAA47315.1; -;
 DR EMBL; AJ010588; CAB82444.1; -;
 DR EMBL; BC014950; AAH14950.1; -;
 DR PIR; I38013; I38013.
 DR HSSP; P00918; I38013.
 DR GeneW; HGNC:1383; CA9.
 DR MIM; 603179; -;
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR GO; GO:0004089; F:carbonate dehydratase activity; TAS.
 DR InterPro; IPR001148; Euk Carbnd.
 DR Pfam; PF000194; carb_anhydrase; 1.
 DR ProDom; PD000865; Euk_CO2_ANHDRASE; 1.
 DR PROSITE; PS00162; Euk_CO2_ANHDRASE; 1.
 DR Lyase; Zinc; Transmembrane; Glycoprotein; Antigen; Signal;
 KW Nuclear protein; Polymorphism.
 FT SIGNAL 1 37
 FT CHAIN 38 459 CARBOIC ANHYDRASE IX.
 FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 251 251 ZINC (CATALYTIC) (BY SIMILARITY).
 FT TRANSMEM 415 435 POTENTIAL.
 FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .).
 FT VARIANT 33 33 M->V.
 FT /FTID=VAR 010787.
 SQ SEQUENCE 459 AA; 49729 MW; 5E3C1E2936050B3F CRC64;

QY 1 MAPLCPSPWPLIPAPAPGLITVOLLSTLLMPVAPORLPKQEDSPFGSSSGEDDL 60
 DB 1 MAPLCPSPWPLIPAPAPGLITVOLLSTLLMPVAPORLPKQEDSPFGSSSGEDDL 60
 QY 61 GEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPQEPQNNARHDXGDQSHMRYGSDPMPWRVSPACAGFQSGVVDIRPOLAFCPLRPL 180
 DB 121 DPQEPQNNARHDXGDQSHMRYGSDPMPWRVSPACAGFQSGVVDIRPOLAFCPLRPL 180
 QY 181 ELLEFQPLPELPLRNNGHSVOLLTPPGLEMAAGREYRALQTLHGAAGRPSSET 240
 DB 181 ELLEFQPLPELPLRNNGHSVOLLTPPGLEMAAGREYRALQTLHGAAGRPSSET 240
 QY 241 VEGHFPFAETHVHLSTAFARVDEALGRPGGLAVLAFLBEGEENSAYEOLLSRLEEA 300
 DB 241 VEGHFPFAETHVHLSTAFARVDEALGRPGGLAVLAFLBEGEENSAYEOLLSRLEEA 300
 QY 301 EEGSETQVPGDLISALLPSDFSSYFQYEGSLTPPCAGQVITVFNQVMTLSAKQLHTLS 360
 DB 301 EEGSETQVPGDLISALLPSDFSSYFQYEGSLTPPCAGQVITVFNQVMTLSAKQLHTLS 360
 QY 361 DTLMGPDSTRLOLNFRAOTPLNGRVLEASPPAGVDSPPAEPVQNSCLAAGDITATVF 420
 DB 361 DTLMGPDSTRLOLNFRAOTPLNGRVLEASPPAGVDSPPAEPVQNSCLAAGDITATVF 420
 QY 421 GLPFAVTSVAFVLYOMRORRGTGQGVSYAPAEVAETGA 459
 DB 421 GLPFAVTSVAFVLYOMRORRGTGQGVSYAPAEVAETGA 459
 DB 421 GLPFAVTSVAFVLYOMRORRGTGQGVSYAPAEVAETGA 459

RESULT 2
 CAH9_MOUSE STANDARD; PRT; 437 AA.
 ID CAH9_MOUSE
 AC Q8VHB5; Q8VHB5; Q8VDE4;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Carbonic anhydrase IX precursor (EC 4.2.1.1) (Carbonate dehydratase IX) (CA-IX) (CAIX) (Membrane antigen MN homolog).
 GN CA9 OR CA9.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=129/Ola;
 RX PubMed=14604546;
 RA Zlatovicova M., Tarabkova K., Syatova E., Gibadulinova A., Mucha V.,
 RA Jakubikova L., Blesova Z., Rafajova M., Ortova Gut M.O., Parkkila S.,
 RA Parkkila A.-K., Wabed A., Sly W.S., Horak I., Pastorek J.,
 RA Pastorekova S.;
 RT "Monoclonal antibodies generated in carbonic anhydrase IX-deficient mice recognize different domains of tumour-associated hypoxia-induced carbonic anhydrase IX";
 RT J. Immunol. Methods 282:117-134(2003).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC Ortova M.;
 RX Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=ICR; TISSUE=Small intestine;
 RA Wang Y.P., Yoshikawa K., Kozaki K., Miyaihi O., Nakagawa A.,
 RA Muratsu H., Kawada Y., Uchida K., Nishikawa N., Saga S.;
 RT "Alternative spliced mRNA coding for MN/CA9";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Reversible hydration of carbon dioxide (By similarity).
 CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
 CC -1- COFACTOR: Zinc (By similarity).

Query Match 100.0%; Score 2424; DB 1; Length 459;
 Best Local Similarity 100.0%; Pred. No. 5.9e-130;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

CC -1- SUBUNIT: Forms oligomers linked by disulfide bonds (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8VHB5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8VHB5-2; Sequence=VSP_007409, VSP_007410;
CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY049077; AAL14193.1; -
DR EMBL; AJ245857; CAC80975.1; -
DR EMBL; AB086322; BAC00816.1; -
DR MGI; 2447188; Car8.
DR InterPro; IPR001148; Euk_Coanhd.
DR Pfam; PF00154; carb_anhydase; 1.
DR ProDom; PD000865; Euk_Coanhd; 1.
DR ProSITE; PS00162; Euk_CO2_ANHYDRASE; 1.
KM Lyase; Zinc; Transmembrane; Glycoprotein; Signal;
KM Alternative splicing.
FT SIGNAL 1 31
FT CHAIN 32 437 CARBONIC ANHYDRASE IX.
FT METAL 205 205 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 207 207 ZINC (CATALYTIC) (BY SIMILARITY).
FT TRANSMEM 230 230 ZINC (CATALYTIC) (BY SIMILARITY).
FT TRANSMEM 391 411 POTENTIAL.
FT CARBOHYD 325 325 N-LINKED (GLCNAC... ) (POTENTIAL).
FT VARSPIC 282 282 G -> V (in isoform 2).
FT VARSPIC 283 437 /FTId=VSP_007409.
FT VARSPIC 283 437 Missing (in isoform 2).
FT FTId=VSP_007410.
SQ SEQUENCE 437 AA; 47264 MW; 88F23380DCD35344 CRC64;
Query Match 65.8%; Score 1596; DB 1; Length 437;
Best Local Similarity 69.4%; Pred. No. 2, 6e-83;
Matches 320; Conservative 36; Mismatches 79; Indels 26; Gaps 6;
QY 1 MAPLCSPMLPILIPAPAGLTVALLLSLMLPVHPOQLPRMGSDFLGGSSGGEEDPL 60
DB 1 MASLGSPWAPLSTPAP---TAQLLLFLLQLQVSAPOQLSGMGQEPFLGDSSSGEDB-L 55
QY 61 GEEDLPSEEDSPREEDPPGEBDLPGEBDLPEVKPKSEEGSKLDELPTVEAPG 120
DB 56 GUVVLPSSEEDAPPEADP-----PGEEDPPEVYSEERMEESGLDLSTPEAP- 102
QY 121 DPQEPONNAHRDQGDQSHRWYGDPPWFRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 103 ---EHSQSHGDKXGGSHWSYGTLLMPQVSPACAGRFQSPVDIRLERTAFORTLOPL 159
QY 181 ELIGPOLPPLRLKLNNGHSVQTLPRGLMVALPGRGYRLQLHLWGAAGPGSEHT 240
DB 160 ELIGYELQPLPELSTLNNGHTVQLTPGLKVALPGRGGYRLQLHLWGSDDPGSEHT 219
QY 241 VEGHRRPAEIHVHLSTAFARVDEALGRGGLAVLAFLLEGPEENSAVEQLRLERLEIA 300
DB 220 VNGHRRPAEIHVHLSTAFARVDEALGRGGLAVLAFLLEGPEENSAVEQLRLERLEIA 279
QY 301 EESSETPVPGDLISALLPSDFSRYYRQESLTPPCACQVITVYNOQVMSAQCLHTLS 360
DB 280 EESSETPVPGDLISALLPSDFSRYYRQESLTPPCACQVITVYNOQVMSAQCLHTLS 339
QY 361 DTLMGQDSRLQINFATQPLNGRVLEAFSPAGVDSPPRAEPVQVNSCLAQDILALVF 420
DB 340 VSLMGQDSRLQINFATQPLNGRVLEAFSPAGVDSPPRAEPVQVNSCLAQDILALVF 396

```

```

QY 421 GILFAVTSVAFINQMRQHR--GTKGVSYSPAETGA 459
DB 397 GILFAVTSVAFINQMRQHR--GTKGVSYSPAETGA 437
RESULT 3
ID CAHE HUMAN STANDARD; PRT; 337 AA.
AC 991X7; GANCF4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Carbonic anhydrase XIV precursor (EC 4.2.1.1) (Carbonate dehydratase
DE XIV) (CA-XIV).
GN CAL4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX SEQUENCE FROM N.A. PubMed=10512682;
RX MEDLINE=99443874; PubMed=10512682;
RA Fujikawa-Adachi K., Nishimori I., Taguchi T., Onishi S.;
RT "Human carbonic anhydrase XIV (CAL4): cDNA cloning, mRNA expression,
RT and mapping to chromosome 1."
RL Genomics 61:74-81(1999).
RN [2]
RP SEQUENCE FROM N.A.
RP Itoget T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto C., Nakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEBO human cDNA sequencing project."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RP TISSUE=Brain, Lung, and Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Martusha K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schreier T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatane P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kerteman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman D.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Butlerfield V.S.N., Krzywinski M.I., Skalska J., Smalins D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Reversible hydration of carbon dioxide.
CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -1- COFACTOR: Zinc (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN ALL PARTS OF THE CENTRAL
CC NERVOUS SYSTEM AND LOWER EXPRESSION IN ADULT LIVER, HEART, SMALL
CC INTESTINE, COLON, KIDNEY, URINARY BLADDER, AND SKELETAL MUSCLE.
CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AB0254765; BAA85002.1; -
 DR EMBL; AC074765; BAC11191.1; -
 DR EMBL; BC034412; AAH34412.1; -
 DR HSSP; O43570; LUD0.
 DR Genem; HGNC:1372; CA14.
 DR MIM; 604832; -
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR GO; GO:0004089; F:carbonate dehydratase activity; TAS.
 DR InterPro; IPR001148; Euk_Coanhd.
 DR Pfam; PF00194; carb_anhydase; 1.
 DR ProDom; PD000865; Euk_Coanhd; 1.
 DR ProSITE; PS00162; Euk_CO2_ANHYDRASE; FALSE_NEG.
 KW Lyase; Zinc; Transmembrane; Signal.
 FT SIGNAL 1 15
 FT CHAIN 16 337
 FT DOMAIN 16 290
 FT TRANSMEM 291 311
 FT DOMAIN 312 337
 FT METAL 109 109
 FT METAL 111 111
 FT METAL 135 135
 FT CARBOHYD 213 213
 FT CONFLICT 229 229
 FT SEQUENCE 337 AA; 37667 MW; 6E101C44BA70A700 CRC64;

Query Match 23.2%; Score 562; DB 1; Length 337;
 Best Local Similarity 38.8%; Pred. No. 3,7e-25;
 Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;

QY 136 DDGSHWRY---GSPRPWRVSPACAGRFQSPVDRPQLAFCAFLPRLGLPQLPRLP 192
 DB 17 DGGHWTYEGPHGQDHMPASTYPCGNNASQSTIDITQTSVTFDDLPALQPHGYDPTGP 76
 QY 193 LRLRNNGHSGVQLTPPGLEMAAGREYRALQLHLHWGAARP-GSEHTVEGHRPAETH 251
 DB 77 LRLNNGHRTVQLSPSTLYLG-GLPKRYVAQLHLHWGQSGSGSEHQSINSPATPAELH 135
 QY 252 VVHL-STAFARVDELDPGGLAVLAFLFEGPEPNNAVYQLSLRELAEGSESGVQVG 310
 DB 136 IVHSDSYDLSLSEAEKPPGLAVGLILEVGETKNAYHILSHLHVKHKKQKTSVFP 195
 QY 311 LDISALPDSFSGRYQYEGSLTTPPCAQGVITVFQNTVMLSAKQHLTSLDTPMGDSR 370
 DB 196 FNLRELPLKQLGQYFRYNGSLTTPPCQSVLMTVFYRSGISWEQLEKLGTLFSTEEBP 255
 QY 371 LQL---NFRATQPLNGRVTENSPFAGVDSFRAEPVQLNSCLAAGDILAVFGLFA-- 425
 DB 256 SKLTVQNYRALQPLNQRFVAFS-----IQAGSYTTGEMSLVGLVGL 302
 QY 426 --VTSVAFVQMERQHRRTGKGVSYRPAE 453
 DB 303 CULLAVYFLARKIRKRLKRNKSVFTFSAQ 332

RESULT 4
 CAHE MOUSE STANDARD; PRT; 337 AA.
 ID CAHE MOUSE STANDARD; PRT; 337 AA.
 AC Q9WV76;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-MAR-2004 (Rel. 43; Last annotation update)
 DE Carboxic anhydrase XIV precursor (EC 4.2.1.1) (Carbonate dehydratase XIV) (CA-XIV).
 GN CA14 OR CAR14 OR CATV.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxId=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=99269110; PubMed=10336468;
 RA Mori K., Ogawa Y., Ebihara K., Tamura N., Teshiro K., Kuwahara T., Mkiyama M., Sugawara A., Ozaki S., Tanaka I., Nakao K.;
 RT "Isolation and characterization of CA XIV, a novel membrane-bound carbonic anhydrase from mouse kidney.";
 RL J. Biol. Chem. 274:15701-15705(1999).

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Stachenko L., Matsushita K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedon T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Locuelliand N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Motley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Muzny D.M., Sodergren E.U., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Rodriguez Y.S.N., Krzywicki M.J., Skalska U., Smalins D.E., Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Reversible hydration of carbon dioxide.
 CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
 CC -1- COFACTOR: Zinc (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Most abundant in the kidney and heart, followed by the skeletal muscle, brain, lung and liver.
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AB005450; BAA78709.1; -
 DR EMBL; BC046985; AAH46995.1; -
 DR HSSP; O43570; LUD0.
 DR WGI; WGI:1344341; Car14.
 DR InterPro; IPR001148; Euk_Coanhd.
 DR Pfam; PF00194; carb_anhydase; 1.
 DR ProDom; PD000865; Euk_Coanhd; 1.
 DR PROSITE; PS00162; Euk_CO2_ANHYDRASE; FALSE_NEG.
 KW Lyase; Zinc; Transmembrane; Signal.
 FT SIGNAL 1 15
 FT CHAIN 16 337
 FT DOMAIN 16 290
 FT TRANSMEM 291 311
 FT DOMAIN 312 337
 FT METAL 109 109
 FT METAL 111 111
 FT METAL 135 135
 FT CARBOHYD 213 213
 FT SEQUENCE 337 AA; 37505 MW; 32F02F4DB78AC0C9 CRC64;

Query Match 23.0%; Score 558; DB 1; Length 337;
 Best Local Similarity 39.4%; Pred. No. 6,2e-25;
 Matches 130; Conservative 42; Mismatches 130; Indels 28; Gaps 8;

QY 136 DQSHWRY---GADPMPRPVSPACAGRFSPVDIRPOLAFCAPALRPLELLGQPLPLE 192
 Db 17 DGHHTVYGRHGHQDHPHSYSPCCGDDQSPINTDTSVITDPDPAVQPHVQLOSTER 76
 QY 193 LKRNHNSVQVLTLPPLGEMALGPREYRALQHLHWAGACR-FGSEHTVEGRHFAEIH 251
 Db 77 LHLNHNHGTVOLSLPPTLHLG-GLPRKYTAQLHLHWQGRSLGSEHQINSEKTAELH 135
 QY 252 VVHL-STAFARVDEALGRPGGLAVLAFLSEEPRENSAYEOLLSLEIMESESETPVPG 310
 Db 136 VHHYSQSYSSLSSEAAQKPGGLAVGLIIEGTEENPAYDILSRHLHRYKQCKSVPP 195
 QY 311 LDLSALLPDSFRYRGYEGSLTTPPCAGVITWYVQNTVMSAQL-----HTLSDTLMGP 366
 Db 196 FSVRELPFQQLQGFRRYNSLTTPPCQSVLMTVFNRAQISMGQLEKLGCTLSSTEDP 255
 QY 367 GDSRLQNFRAIQEINGRVIETASFAGVDSFRAEPYQNSCLAAGDILALVGLLFA- 425
 Db 256 SEPLVQ-NYRVPQPLNQRITFASF-----IQAGPLYTGBMLGLGVGILAGC 301
 QY 426 ---VTSVAFVQMRROHRRGRGTGVSYPFA 452
 Db 302 LCILLAVYFLAQKIRKRLGNKSVVPTSA 331

RESULT 5

CAHC_HUMAN STANDARD; PRT; 354 AA.
 AC 043570; Q9BEG2;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Carbonic anhydrase XII precursor (EC 4.2.1.1) (Carbonate dehydratase
 XII) (CA-XII) (Tumor antigen HOM-RCC-3.1.3).
 GN CA12.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
 RC TISSUE=Renal cell carcinoma;
 RX MEDLINE=98301622; PubMed=9636197;
 RA Tuercoi C., Sahin U., Vollmar E., Siemer S., Goertert E., Seitz G.,
 RT "Human carbonic anhydrase XII: cDNA cloning, expression, and
 RT chromosomal localization of a carbonic anhydrase gene that is
 RT overexpressed in some renal cell cancers."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:7608-7613(1998).
 RN (2)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Lung;
 RX MEDLINE=98445416; PubMed=9770531;
 RA Ivanov S.V., Kuzmin I., Wei M.-H., Pack S., Geil L., Johnson B.E.,
 RT "Down-regulation of transmembrane carbonic anhydrases in renal cell
 RT carcinoma cell lines by wild-type von Hippel-Lindau transgenes."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12596-12601(1998).
 RN (3)
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Eye, and Kidney;
 RX MEDLINE=22388257; PubMed=1247932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares W.B., Bonaldo M.F., Casavant T.P., Scheetz T.E.,
 RA Brownstein M.J., Useth T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muljany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.V., Hulik S.M.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (4)
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 30-292.
 RX MEDLINE=21396545; PubMed=11493685;
 RA Whittington D.A., Waheed A., Ulmasov B., Shah G.N., Grubb J.H.,
 RA Sly W.S., Christianson D.W.,
 RT "Crystal structure of the dimeric extracellular domain of human
 RT carbonic anhydrase XII, a bitoric membrane protein overexpressed in
 RT certain cancer tumor cells."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9545-9550(2001).
 CC -1- FUNCTION: Reversible hydration of carbon dioxide.
 CC -1- CATALYTIC ACTIVITY: H(2)O + CO(2) = HCO(3) + H(2)O.
 CC -1- COFACTOR: Zinc (by similarity).
 CC -1- ENZYME REGULATION: Inhibited by acetazolamide.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=O43570-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O43570-2; Sequence=VSP_000772;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN COLON, KIDNEY, PROSTATE,
 CC INTESTINE AND ACTIVATED LYMPHOCYTES. EXPRESSED AT MUCH HIGHER
 CC LEVELS IN THE RENAL CELL CANCERS THAN IN SURROUNDING NORMAL KIDNEY
 CC TISSUE. MODERATELY EXPRESSED IN PANCREAS, OVARY, AND TESTIS.
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 CC EMBL, AF051862; AAC39789.1; -
 CC EMBL, AF053735; AAC63952.1; -
 CC EMBL, BC000278; AAH00278.1; -
 CC EMBL, BC011691; AAH11691.1; -
 CC EMBL, BC023981; AAH23981.1; -
 CC PDB, 1JDO; 17-AUG-01.
 CC PDB, 1JCZ; 17-AUG-01.
 CC GeneW, HSMC:1371; CA12.
 CC MIM, 603263; -
 CC GO, GO:0016021; C:integral to membrane; TAS.
 CC GO, GO:0004089; F:carbonate dehydratase activity; TAS.
 CC GO, GO:0008270; F:zinc ion binding; TAS.
 CC InterPro, IPR001148; Bsk_Coanhd.
 CC Pfam, PF00194; carb_anhydrase.1.
 CC ProDom, PD000865; Bsk_Coanhd.1.
 CC PROSITE, PS00162; EUK_CO2_ANHYDRASE; 1.
 CC Lysase; Zinc; Transmembrane; Signal; Alternative splicing;
 CC 3D-structure.
 CC SIGNAL 1 24
 CC CHAIN 25 354
 CC DOMAIN 25 301
 CC TRANSMEM 302 322
 CC DOMAIN 323 354
 CC METAL 119 119
 CC METAL 121 121
 CC METAL 145 145
 CC DISULFID 50 230
 CC POTENTIAL.
 CC CARBONIC ANHYDRASE XII.
 CC EXTRACELLULAR (POTENTIAL).
 CC POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).
 CC ZINC (CATALYTIC).
 CC ZINC (CATALYTIC).
 CC ZINC (CATALYTIC).
 CC ZINC (CATALYTIC).

```

FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT VARSPLIC 292 302 Missing (in isoform 2).
/ftid=vsp_000772.

FT TURN 37 38
FT HELIX 40 42
FT HELIX 43 46
FT HELIX 48 51
FT STRAND 59 60
FT HELIX 62 64
FT STRAND 65 67
FT TURN 69 70
FT STRAND 75 77
FT STRAND 80 80
FT TURN 83 84
FT STRAND 86 91
FT STRAND 96 99
FT TURN 102 103
FT STRAND 105 107
FT STRAND 113 122
FT STRAND 124 124
FT TURN 125 126
FT STRAND 127 127
FT STRAND 134 135
FT TURN 136 137
FT STRAND 138 138
FT STRAND 142 150
FT TURN 151 153
FT HELIX 157 160
FT TURN 161 162
FT TURN 164 165
FT STRAND 167 176
FT HELIX 181 187
FT TURN 188 189
FT HELIX 190 193
FT TURN 196 197
FT STRAND 199 203
FT STRAND 206 206
FT HELIX 207 210
FT TURN 215 216
FT STRAND 218 223
FT TURN 228 229
FT STRAND 234 239
FT STRAND 243 245
FT HELIX 247 255
FT STRAND 258 258
FT TURN 262 263
FT STRAND 269 269
FT STRAND 285 287
SQ SEQUENCE 354 AA, 39451 MW, 9016216B2C6C0C CRC64;

Query Match 22.2%; Score 537.5; DB 1; Length 354;
Best Local Similarity 35.8%; Pred. No. 9.4e-24;
Matches 129; Conservative 46; Mismatches 142; Indels 43; Gaps 10;

```

```

DB 246 SQEOLALLETALCTHMDPSPREMINNROVQKDERLVTYSF-----QV 292
QY 407 NSCIAAGDIALVFGLLFA-----VTSVAFVQMRQHRGTGVSYPRAEVAETGA 459
DB 293 QVCTAAGLSIGILSLALAGILGICIVVVVSIWLFRRKIKKDKNGVITYKPMETEA 352

RESULT 6
CAH6_SHEEP STANDARD; PRT; 307 AA.
AC P08060;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Carbonic anhydrase VI (EC 4.2.1.1) (Carbonate dehydratase VI) (CA-VI)
DE (Secreted carbonic anhydrase) (Salivary carbonic anhydrase).
GN CA6.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RC TISSUE=Saliva;
RX MEDLINE=68294021; PubMed=315834;
RA Fernley R.T., Wright R.D., Coghlan J.P.;
RT "Complete amino acid sequence of ovine salivary carbonic anhydrase.";
RL Biochemistry 27:2815-2820(1988).
CC -1- FUNCTION: Reversible hydration of carbon dioxide. Its role in
CC saliva is unknown.
CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -1- COFACTOR: Zinc (by similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Major constituent of saliva.
CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC family.
DR PIR: A29993; A29993.
DR HSSP: O43570; IUD0.
DR InterPro: IPR001148; Euk Coanhd.
DR Pfam: PF00194; carb anhydrase; 1.
DR ProDom: PD000865; Euk Coanhd; 1.
DR PROSITE: PS00162; Euk CO2 ANHYDRASE; 1.
KW lyase; Zinc; Glycoprotein.
FT METAL 94 94 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 96 96 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 121 121 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 25 207 POTENTIAL.
FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .).
FT VARIANT 63 63 V -> M.
FT VARIANT 297 297 I -> M.
SQ SEQUENCE 307 AA, 35555 MW, 3368822D45B5D6C CRC64;

Query Match 21.7%; Score 526; DB 1; Length 307;
Best Local Similarity 43.2%; Pred. No. 3.6e-23;
Matches 111; Conservative 40; Mismatches 98; Indels 8; Gaps 4;

```



```

CC -----
DR EMBL: X96503; CA65357.1;
DR PIR: S71877; S71877.
DR HSRP: Q43570; IJ00.
DR InterPro: IPR001148; Euk Coanhd.
DR Pfam: PF00194; carb_anhydase; 1.
DR ProDom: PD000865; Euk Coanhd; 1.
DR ProSITE: PS00162; Euk CO2 ANHYDRASE; 1.
DR Lyase; Zinc; Glycoprotein; Signal.
FT SIGNAL 1 14
FT CHAIN 1 319
FT METAL 106 319 CARBONIC ANHYDRASE VI.
FT METAL 106 106 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 108 108 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 133 133 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 37 219 POTENTIAL.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .).
FT CONFICT 16 16 H -> S (IN REF. 2).
SQ SEQUENCE 319 AA; 37006 MW; 17984847A9083AED CRC64;

Query Match 21.1%; Score 511; DB 1; Length 319;
Best Local Similarity 42.0%; Pred. No. 2.6e-22;
Matches 108; Conservative 40; Mismatches 97; Indels 12; Gaps 4;

QY 137 DQSHWRYGDDPPWPRVSPACAGRFSPVDIPQLAAFCFPAIRPELLGFQLPPELR 196
DB 26 DEKRWRL-----QYDCCGGRSPFDLKKKKRVNPSRLNLTGYGURQ-GEPPMT 76
QY 197 NNGHSYQTLTPGLLEALGPGRHYALQLHLHWG--AAGRGSHTEGHRPAEIVHV 254
DB 77 NNGHTYQILSPSSMNTSDSGYAKGMHPHMGDSSEISGSHYDGMKRYIIEHVH 136
QY 255 LSTAFARVDALGRPGGLAVLAFLF-EGPEBNAYEQLSLLEIEEGSEFQVPELDI 313
DB 137 YHSKXGSEAEQNEDEGLAVLAALVEVDYENTYNSFIHLDIYVAGQSTYLRDLDI 196
QY 314 SALPSPDFSRFYQYEGSLTPPCAGVITWENQYVWLSAQQLHLSDTLWPGDSRLQL 373
DB 197 QDMPLDGLRYVSYLSGLTTPSCNTENHMFVADIVKLSKQIEKLSNLTNOMETION 256

QY 374 NFRATQPLNGRYVIEASF 390
DB 257 NYRSTQPLNRYVEANF 273

RESULT 9
CAHC_MOUSE STANDARD; PRT; 354 AA.
AC 08C185; 08BK6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Carbonic anhydrase XII precursor (EC 4.2.1.1) (Carbonate dehydratase
DE XII) (CA-XII).
GN CA12 OR CAR12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
OX NCBI_TaxID=10090;
RN 1[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi T., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bull T.P., Hume D.A., Quackenbush J.,
RA Schirni L.M., Kanapin A., Matluda H., Baralov S., Beisel K.W.,
RA Blake U.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedziarski R.M., King B.L.,

```

```

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagasima T., Numata K., Okido T., Pavan W.J., Pereira G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius U.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Walestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Zimner P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawat J., Aizawa K., Arikawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishi Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino K., Yashiro R., Yashiro R., Yashiro R.,
RA Yoney E., Hayashizaki Y.
RA Birtney E., Hayashizaki Y.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopfins R.F., Jordan H., Moore T., Max S.T., Wang D., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Tohiyuki S., Carninci P., Prange C.,
RA Rabb S.S., Logguelano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Patey J., Helton E., Kettman W., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boufield G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalins D.E.,
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16703(2002).
CC - FUNCTION: Reversible hydration of carbon dioxide.
CC - CATALYTIC ACTIVITY: H(2)O(3) = CO(2) + H(2)O.
CC - COFACTOR: zinc (By similarity).
CC - SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC - SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AK052639; BAC35074.1;
DR EMBL: BC035941; AAH35941.1;
DR InterPro: IPR001148; Euk Coanhd.
DR Pfam: PF00194; carb_anhydase; 1.
DR ProDom: PD000865; Euk Coanhd; 1.
DR ProSITE: PS00162; Euk CO2 ANHYDRASE; UNKNOWN_1.
LYase; Zinc; Transmembrane; Signal.
FT SIGNAL 1 24
FT CHAIN 25 354
FT DOMAIN 25 301 CARBONIC ANHYDRASE XII.
FT TRANSMEM 302 322 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 323 354 POTENTIAL.
FT METAL 120 120 CYTOPLASMIC (POTENTIAL).
FT METAL 122 120 ZINC (CATALYTIC).
FT METAL 146 146 ZINC (CATALYTIC).

```

FT DISULFID 50 221 BY SIMILARITY.
 FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 297 297 D -> N (IN REF. 1).
 SQ SEQUENCE 354 AA; 39695 MW; 98B581BF2A1611 CRC64;
 Query Match 20.5%; Score 496; DB 1; Length 354;
 Best Local Similarity 35.5%; Pred. No. 2,1e-21;
 Matches 115; Conservative 54; Mismatches 137; Indels 18; Gaps 7;
 QY 139 SHWRY---GDPMPRVSPACAGRFQSPVDIRPOLAFCPALRELLGQLPPLPELRL 195
 DB 30 SKWTYVGPAGEKXWKSXKYPSCGGLQSPIDHSIDLQYDASIALPQGYVWVSXEKLNLT 89
 QY 196 RNNGHSVQVLTLPGLGEMALGPGRVETALQHLHNGAAGR--GSEHTYEGHFRPALTIVH 254
 DB 90 TNDGSHVRLNLSMDVYIQLOPHYRAQLHLHNGNRNDPQSGHVSQGHFAELHIVH 149
 QY 255 L-STAFARVDEALGRPGGLAVLAFLBEGPENSAVEQLSLREIEAEGSETOVPLDI 313
 DB 150 YNSDLYPDSTASDSEGLAVLAVLIEIG--SAPSYDKIFSHLGHVYKQGVLIPOFNI 208
 QY 314 SALLPSDSRYPQYQSGSLTTPPCAGGVWTFVFNQTMLSAKQHTLSDTLM-----GPGD 368
 DB 209 EELLPSPEGEYRYRGSLTTPPCVPTVMTVFRNPVQISOQLALSTALYFTHMDDPTP 268
 QY 369 SRLQNFRAQPLNGRVIEASFPAGVDSPPAAEPVQINSLAAGDILALVFGLLPAVTS 428
 DB 269 REMINPFQVQKFDRLVYISFROGLLDTGLSLGILLSVALA-----VLGIISI-VLA 321
 QY 429 VAFIVQMRQHRRTKGVSTYRPA 452
 DB 322 VSIWLFKRKKSKKGNKVYKPA 345
 RESULT 10
 ID CAH6_HUMAN STANDARD; PRT; 308 AA.
 AC P23280; Q96QX8; Q9UF03;
 DT 01-NOV-1991 (Rel. 20; Created)
 DT 28-FEB-2003 (Rel. 41; Last sequence update)
 DT 10-OCT-2003 (Rel. 42; Last annotation update)
 DE Carbonic anhydrase VI precursor (EC 4.2.1.1) (Carbonate dehydratase VI) (CA-VI) (secreted carbonic anhydrase) (salivary carbonic anhydrase).
 GN CA6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=51105141; PubMed=1699030;
 RX Aided P., Fu P., Barrett G., Penschow J.D., Wright R., Coghlan J.P., Fernley R.T.;
 RA "Human secreted carbonic anhydrase: cDNA cloning, nucleotide sequence, and hybridization histochemistry.";
 RT Biochemistry 30:569-575(1991).
 RN [2] SEQUENCE FROM N.A.
 RP Grubb D.J.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3] SEQUENCE FROM N.A.
 RA Frankland J.;
 RP Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Reversible hydration of carbon dioxide. Its role in saliva is unknown.
 CC -1- CATALYTIC ACTIVITY: H(2)O + CO(2) -> H(2)O.
 CC -1- COFACTOR: Zinc (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Major constituent of saliva.
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M57692; AAF51892.1; -
 DR EMBL; AF128418; AAF22565.1; JOINED.
 DR EMBL; AF128411; AAF22565.1; JOINED.
 DR EMBL; AF128412; AAF22565.1; JOINED.
 DR EMBL; AF128413; AAF22565.1; JOINED.
 DR EMBL; AF128414; AAF22565.1; JOINED.
 DR EMBL; AF128415; AAF22565.1; JOINED.
 DR EMBL; AF128416; AAF22565.1; JOINED.
 DR EMBL; AF128417; AAF22565.1; JOINED.
 DR HSSP; AL139415; CAC42429.1; -
 DR HSSP; O43570; IUD0.
 DR GeneW; HGNC:1380; CA6.
 DR MIM; 144780; -
 DR GO; GO:0004089; P:carbonate dehydratase activity; TAS.
 DR InterPro; IPR001148; Euk Coahd.
 DR Pfam; PF00194; carb_anhydrase; 1.
 DR ProDom; PD000865; Euk Coahd; 1.
 DR ProSite; PS00162; Euk CO2 ANHYDRASE; 1.
 DR KX; KX: -
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 308 CARBONIC ANHYDRASE VI.
 FT METAL 111 111 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 113 113 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 138 138 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 42 224 POTENTIAL.
 FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 90 90 G -> S (IN REF. 3).
 FT CONFLICT 103 103 T -> I (IN REF. 1).
 FT CONFLICT 148 148 T -> S (IN REF. 3).
 FT CONFLICT 270 270 K -> N (IN REF. 3).
 SQ SEQUENCE 308 AA; 35365 MW; 5B6BFF2DD239333 CRC64;
 Query Match 20.1%; Score 488; DB 1; Length 308;
 Best Local Similarity 39.2%; Pred. No. 5e-21;
 Matches 102; Conservative 44; Mismatches 106; Indels 8; Gaps 4;
 QY 139 SHWRYG---GDPMPRVSPACAGRFQSPVDIRPOLAFCPALRELLGQLPPLPELRL 194
 DB 21 SDWYSGALDEAMWPHYPAQGGQSGSPINLGRTVRVNPLKGLNMTGYE-TQAGEFP 79
 QY 195 LRNGHSVQVLTLPGLGEMALGPGRVETALQHLHNGAAGR--GSEHTYEGHFRPALTIV 252
 DB 80 WNNGHVTOIGLPTMTMTYADGTIVYIAQMFHMGASSEISGSHTVADGRLHVIH 139
 QY 253 VHLSTAFARVDEALGRPGGLAVLAFLBEGPENSAVEQLSLREIEAEGSETOVPLDI 311
 DB 140 VHNYSKYTDIADADAGLAVLAFLBEGPENSAVEQLSLREIEAEGSETOVPLDI 199
 QY 312 DISALPSDSRYPQYQSGSLTTPPCAGGVWTFVFNQTMLSAKQHTLSDTLMGPDSDL 371
 DB 200 DVQMLPEPRNQHYTYSGLTTPPCETENHFWLADQVXLSRTQVKNLSLDRNKTI 259
 QY 372 QLNFRATQPLNGRVIEASFP 391
 DB 260 HNDYRRTQPLKRVVSNFP 279
 RESULT 11
 ID CAH6_MOUSE STANDARD; PRT; 317 AA.

AC P18761: 088625; 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Carbonic anhydrase VI precursor (EC 4.2.1.1) (Carbonate dehydratase
 VI) (CA-VI) (Secreted carbonic anhydrase) (Salivary carbonic
 anhydrase).
 DE (CA-VI) (Secreted carbonic anhydrase) (Salivary carbonic
 anhydrase).
 CN CA6 OR CAR6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Salivary gland;
 RX MEDLINE=99077987; PubMed=9856573;
 RA Sok J., Wang X.Z., Batchvarova N., Kuroda M., Harding H., Ron D.,
 RT "Chop-dependent stress-inducible expression of a novel form of
 RT carbonic anhydrase VI.";
 RL Mol. Cell. Biol. 19:495-504(1999).
 RN [2]
 RP SEQUENCE OF 18-38.
 RX MEDLINE=89246331; PubMed=2497732;
 RA Fenley R.T., Darling P., Aldred P., Wright R.D., Coghlan J.P.,
 RT "Tissue and species distribution of the secreted carbonic anhydrase
 RT isoenzyme.";
 RL Biochem. J. 259:91-96(1989).
 CC -1- FUNCTION: Reversible hydration of carbon dioxide. Its role in
 CC saliva is unknown.
 CC -1- CATALYTIC ACTIVITY: H(2)O(3) = CO(2) + H(2)O.
 CC -1- COFACTOR: Zinc (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Major constituent of saliva.
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
 CC family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF079835; AAD1540.1; -
 DR PIR: S03863; S03863.
 DR HSP: P00918; IBV3.
 DR MGI: MGI:1333786; Car6.
 DR InterPro: IPR001148; Euk_Coand.
 DR Pfam: PF00194; carb_anhydrase; 1.
 DR ProDom: PD000865; Euk_Coand; 1.
 DR PROSITE: PS00162; Euk_CO2_ANHYDRASE; FALSE_NEG.
 KM Lyase; Zinc; Glycoprotein; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 317 CARBONIC ANHYDRASE VI.
 FT METAL 110 110 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 112 112 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 137 137 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 41 223 POTENTIAL.
 FT CARBOHYD 255 255 N-LINKED (GLYCAC. .) (POTENTIAL).
 FT CONFLICT 18 18 H -> G (IN REF. 2).
 FT CONFLICT 35 35 S -> T (IN REF. 2).
 FT CONFLICT 35 35 S -> T (IN REF. 2).
 SQ SEQUENCE 317 AA; 36348 MW; 0393844F7B48283A CRC64;
 Query Match 19.3%; Score 467.5; DB 1; Length 317;
 Best Local Similarity 37.4%; Pred. No. 7.4e-20;
 Matches 98; Conservative 45; Mismatches 106; Indels 13; Gaps 5;

DB 78 TWNNNGHTYIDLPSPMYLETSDGTEP.SKAFFHMG--GRDWELSGEHTIDGIRSI 135
 QY 250 IYVHLSSTFAVDALGRPGGLAVIAA-FLEEGSENAVEQLSREELAESENOV 308
 DB 136 AHFVHNKRYGTYEENKODKGNLAVLAFKIDFAEYVSDIISALDKERGETTL 195
 QY 309 PGIDISALLPDSFSEYFQYEGSLTPPCAGVITVFNQYMLSAKQHLTSLTLMGPD 368
 DB 196 KDTINDLDPKDVHNYHTYPSGLTTPCTENVQVFLDRVTLNQAQVYTIENVMDDNN 255
 QY 369 SRLQNFRAQTPQPLNGVITASF 390
 DB 256 NTIQNGYRSTQPNHRYVEANF 277
 RESULT 12
 ID CAHF_MOUSE STANDARD; PRT; 324 AA.
 AC Q99N23;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Carbonic anhydrase XV precursor (EC 4.2.1.1) (Carbonate dehydratase
 DE XV) (CA-XV).
 GN CA15 OR CAR15.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RA Hewett-Emmett D., Shumlin L.C.,
 RT "Characterization and evolution of two new members of the alpha-
 RT carbonic anhydrase gene family in mouse: Car13 and Car15.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=2386257; PubMed=12477932;
 RA Strauberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marxson K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein W.J., Uedin T.B., Toshlyuk S., Cantoni P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Reversible hydration of carbon dioxide (By similarity).
 CC -1- CATALYTIC ACTIVITY: H(2)O(3) = CO(2) + H(2)O.
 CC -1- COFACTOR: Zinc (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
 CC family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL, AF231122; AAK16671.1; -

CC EMBL, BC019975; AAK19975.1; -

DR HSSP; G64444; 22NC.

DR MGD; MG1:1931324; Car15.

DR InterPro; IPR001148; Euk Coand.

DR Pfam; PF00194; carb anhydrase; 1.

DR ProDom; PD000665; Euk Coand; 1.

DR PROSITE; PS00162; Euk_CO2_ANGHDASE; FALSE_NEG.

KW Lyase; Zinc; Glycoprotein; Signal.

FT SIGNAL 1

FT CHAIN 19

FT METAL 122

FT METAL 124

FT METAL 147

FT CARBOHYD 184

FT CARBOHYD 194

FT CARBOHYD 203

SEQUENCE 324 AA; 35482 MW; D285DD086476372F CRC64;

Query Match 17.4%; Score 422.5; DB 1; Length 324; Best Local Similarity 35.3%; Pred. No. 2.6e-17; Matches 107; Conservative 42; Mismatches 125; Indels 29; Gaps 8;

QY 136 DDQSHWRYGDP-----WPRVSPACAGRFQSPVDIRPOLAFCPLRPLELGFQLPP 189

DB 20 DSSGTWCYSDQPKCGPAHWKELAPACGPPSPINIDLRLVQDRDYTLKEFFIRQGYDSAP 79

QY 190 LPELRLRNHNSVQITLP-----PGLMALGPRREVRALQILHMGAAARPGSEHTVEG 243

DB 80 QDPWVLENDGHTVLIRVNSCQNCPLAIGAGLPSPEYKLLQLHFMWSPGHQSEHSLDE 139

QY 244 HRPFAIHVHLSTAFARVDEALGRPGILAVLAFLIEGPEENSAVYQLSRLEIAEES 303

DB 140 KKGSMEMHVNHTNYQSWEDARSPDGFALLAVLVEBDNNTNFSAIVSGKLNSSPG 199

QY 304 SETQVPG-LDIALLPSPDS--RFQYFGSLTPPCAGQVWTFVFNQTVLMSAKQHTLS 360

DB 200 VAVNLTSTPALASTLPSALRLRLRYRYSGLTTPCEBPVAVMTVFENTVPIGLAQVQFO 259

QY 361 DTL-WQPGDSR---IQNFRATQPLNGRVIENAFPGVDSPPRAEPVQNSCIAADIL 416

DB 260 AVLGCPGGLHPRPLTNSRNPQPLGGRISAPSAVRS-----VSTLPGLH-----L 309

QY 417 ALV 419

DB 310 ALV 312

RESULT 13

CAH2_TRIHK STANDARD; PRT; 259 AA.

ID CAH2_TRIHK

AC 08UMAS;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Carbonic anhydrase II (BC 4.2.1.1) (carbonate dehydratase II) (CA-II).

GN CA2.

OS Tribolodon hakonensis (Japanese dace).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Tribolodon.

OC NCBI_Taxid=151740;

OX NCBI_Taxid=151740;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22562542; PubMed=12531781;

RA Hirata T., Kaneko T., Ono T., Nakazato T., Furukawa N., Hasegawa S., Wakabayashi S., Shigekawa M., Chang M.H., Romero M.F., Hirose S., Am. J. Physiol. 284:R1199-R1212(2003).

RT "Mechanism of acid adaptation of a fish living in a pH 3.5 lake."

RL FEBS Lett. 271:137-140(1990).

CC -!- FUNCTION: Reversible hydration of carbon dioxide.

CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.

CC -!- COFACTOR: Zinc (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL; AB055617; BAB83090.1; -

DR InterPro; IPR001148; Euk Coand.

DR Pfam; PF00194; carb anhydrase; 1.

DR ProDom; PD000665; Euk Coand; 1.

DR PROSITE; PS00162; Euk_CO2_ANGHDASE; 1.

KW Lyase; Zinc.

FT INIT_MET 0

FT METAL 93

FT METAL 95

FT METAL 118

SEQUENCE 259 AA; 29002 MW; 622D13D1C2378B86 CRC64;

Query Match 17.4%; Score 422; DB 1; Length 259; Best Local Similarity 36.9%; Pred. No. 2.2e-17; Matches 97; Conservative 46; Mismatches 104; Indels 16; Gaps 8;

QY 139 SH-WRY---GGDPWPRVSPACAGRFQSPVDIRPOLAFCPLRPLELGFQLPP 194

DB 1 SHGCGYADHNGPQKCNENFPIANGPROSPIDITQTKASVYDTIKPKL---KVDPTTSD 57

QY 195 LRNGHSVQVLTLPGLMAL---GP-GREYRALQLHMGAAARPGSEHTVEGHRPAEI 250

DB 58 ILNNGHSFQVTFADNDSSMLTEGPISGKRLQLQHFHMGASGKSGSEHTVDCXPAEL 117

QY 251 HVVHSTAFARVDEALGRPGILAVLAFLIEGPEENSAVYQLSRLEIAEESSETQVPG 310

DB 118 HLVMNNTKIASFGAPANKPGAVGVFLDIG-EDNPKLQKIDAMDATSKKQKQSF 176

QY 311 LDIALLPSPDSRFQYFGSLTPPCAGQVWTFVFNQTVLMSAKQHTLSDTLWPGDSR 370

DB 177 FDPCTLPKSP-EWYTFGSLTPPLEYESTVWIVCKQPIVSSEQMKKFRSLFTAEER 235

QY 371 ---IQNFRATQPLNGRVIENAF 390

DB 236 ACCMVNNYRPPQPLKDKKVCASF 258

RESULT 14

CAH8_MOUSE STANDARD; PRT; 290 AA.

ID CAH8_MOUSE

AC P28651; O8GCF5; O91XF6;

DT 01-DEC-1992 (Rel. 24, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Carbonic anhydrase-related protein (CARP) (CA-VIII).

GN CA8 OR CAR8 OR CALSI OR CALS OR CARP.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_Taxid=10090;

OX NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=BALB/c; TISSUE=Brain; MEDLINE=2121526;

RA Kato K.;

RT "Sequence of a novel carbonic anhydrase-related polypeptide and its exclusive presence in Furukawa cells."

RL FEBS Lett. 271:137-140(1990).

CC [2]

RP SEQUENCE FROM N.A., SEQUENCE OF N-TERMINUS, AND MUTAGENESIS.
 RX MEDLINE=92131620; PubMed=8977131;
 RA Stedman B., Elledy B., Wallgren K., Jonsson B.-H., Lindskog S.;
 RT Two point mutations convert a catalytically inactive carbonic
 RL anhydrase-related protein (CARP) to an active enzyme.";
 RL FEBS Lett. 398:322-325(1996).
 [3]
 RP SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=22354683; PubMed=1246651;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nakada I., Osato N., Saito R., Suzuki H., Yamataka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matheida H., Batilov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbin L.E., Cousins S.,
 RA Datta E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gasteier T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Gilmond S., Gustinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kana A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan M.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Portius J.U., Qi D., Ramchandran S.,
 RA Ravasi T., Reed U.C., Reed D.U., Reid J., Ring B.Z., Ringwald M.,
 RA Sanchelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilmng L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Zimnick P., Hayatsu N.,
 RA Hiranano-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imocant K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyatake A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Struhsberg R.L., Feinsold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhac N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman A.C., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Does not have a carbonic anhydrase catalytic activity.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN PURKINE CELLS.
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
 CC family.

CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: X6197; CNA4368.1; -
 DR EMBL: AK004886; BAC25100.1; -
 DR EMBL: BC010773; AAH10773.1; -
 DR PIR: S12867; S12867.
 DR HSP: P00918; IRY3.
 DR MGD: MGI:88253; Car8.
 DR InterPro: IPR001148; Euk Coanhd.
 DR Pfam: PF00194; carb anhydrase.1.
 DR PROSITE: PS00162; ETK_CO2_ANTIHYDASE.1.
 KW Zinc.
 FT INIT MET 0
 FT DOMAIN 0
 FT SITE 14 36 GLU-RICH (ACIDIC).
 FT SITE 116 116 ANCESTRAL ZINC LIGAND.
 FT METAL 118 118 ZINC (POTENTIAL).
 FT METAL 141 141 ZINC (POTENTIAL).
 FT METAL 141 141 ZINC (POTENTIAL).
 FT METAL 114 116 EVR->OVR: RESTORE ZINC-BINDING AND
 FT ACTIVITY
 FT CONFLICT 260 260 A->V (IN REF. 1).
 FT CONFLICT 284 284 V->I (IN REF. 3).
 FT CONFLICT 290 290 Q->S (ORRECTED) (IN REF. 1).
 SQ SEQUENCE 290 AA; 32950 MW; D0A998A335C8E502 CRC64;
 Query Match 17.3%; Score 419.5; DB 1; Length 290;
 Best Local Similarity 33.3%; Pred No. 3.4e-17;
 Matches 99; Conservative 59; Mismatches 112; Indels 27; Gaps 9;
 QY 112 DLPLVE-APSDPGEPPONNAHREKGDQSHWRYGDPMPRVSFACAGFQSPVDIRPOL 170
 DB 2 DLSDIEDVAVFPKEEEDDEEEVE---WGEEGVGLVFPDNGEYQSPIMNSRE 58
 QY 171 AAFPPARPELIGFQPP-----LPRLRNNSVQLPLPGLEMAAGP---GEYRL 223
 DB 59 ARVPS-----LDVRLSPVTVCRCEVYNDHTIQLVLSKSVSGGLPQGEFELY 113
 QY 224 QHLHWGAGRPSSEHTVEGRPAEIVHVL-STAFARVDEALRPGGLAVIAFLBERG 282
 DB 114 EVRFHWGRNQRESEHTVNFKAPPELHLIHNWSTLFGSIDAVGPHDIALIAFLVQIG 173
 QY 283 PEENSAVEQLSLREIEAGSETOVPGDISALPDSR-YFOYEGSLTPPCAQYI 341
 DB 174 -KEVVGKAVTEILQDIQYKSKTICGNEPTLLDPPLRQVWYEGSLTTPSEGV 232
 QY 342 WYFNQTVMLSAQL-----HTLSDTLWGPGRSRLQINFRATQPINGRYTEASF 390
 DB 233 WILFRYPLTISQVIEEPRRLRTHVKGALVSGCGIIGDNRPTQPSIDRVIRAAF 289
 RESULT 15
 CAH8_HUMAN STANDARD; PRT; 289 AA.
 ID CAH8_HUMAN
 AC P35219;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Carbonic anhydrase-related protein (CARP) (CA-VIIT).
 GN CA8 OR CALS.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246262; PubMed=8462548;
 RA Skaggs L.A., Bergenhien N.C.H., Venta P.J., Tashian R.E.;
 RT "The deduced amino acid sequence of human carbonic anhydrase-related
 RL protein (CARP) is 98% identical to the mouse homologue".
 CC -1- FUNCTION: Does not have a carbonic anhydrase catalytic activity.
 CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
 CC family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L04656; AAA35653.2; -
DR PIR; JN0576; JN0576.
DR HSSP; P00918; 1B73.
DR Genew; HGNC:1382; CAB.
DR MIM; 114815; -
DR CO; GO:0004088; F:carbonate dehydratase activity; TAS.
DR InterPro; IPR001148; Euk_Coamd.
DR Pfam; PF00194; carb_anhydrase; 1.
DR ProDom; PD000865; Euk_Coamd; 1.
DR PROSITE; PS00162; Euk_CO2_ANHYDRASE; 1.
KM Zinc.
FT INIT MET 0 0 BY SIMILARITY.
FT DOMA TN 14 35 GLU-RICH (ACIDIC).
FT SITE 115 115 ANCESTRAL ZINC LIGAND.
FT METAL 117 117 ZINC (POTENTIAL).
FT METAL 140 140 ZINC (POTENTIAL).
SQ SEQUENCE 289 AA; 32842 MW; BA6B5A932A324BE CRC64;

Query Match: 17.0%; Score 412; DB 1; Length 289;
Best Local Similarity: 33.6%; Pred. No. 9e-17;
Matches 100; Conservative 56; Mismatches 110; Indels 32; Gaps 10;

QY 110 LEDLPTVAPGDDPQEPQNNARHDKESDDQSHWKYGGDDPWPRVSPACAGRFQSPYDIRQ 163
DB 6 LED--TVAFP--EKEDEEEEEEVE--WGIEGVEWGLVFPDANGELYQSPINLSR 56
QY 170 LAAPCAPLPLELIGFCPP--LPELRNNGHSYQTLTPGLEMALGP--GREYRA 222
DB 57 EARVDP-----LDVRLSPNYVYVCRDCEVTNDGHTIQVILKSKSVLSGGPLQGHPEL 111
QY 223 LQHLHWGAGAPGSGSEHTVEGHRPAEIHVHL-STFAAVDEALGRPGGLAVLAFLBE 281
DB 112 YVRFHWRGRENQRGSEHTVNFKAPMEHLIHMNSTLFGSIDAVGKPGHAIIALFVQI 171
QY 282 GPENSAYEQLLSRLAEIAGESETOVPGDISALPDSFR--YFQYEGSLTTPCAQGV 340
DB 172 G-KHVGAKAVTEIQDIQYKSKXTPCENPTLLPDLRLDYVWYEGSLTTPCSEGV 230
QY 341 TWTVFNQVWMLSAKL-----HTLSDTLWGPDSRLQANFRATQPLNGRYTEASF 390
DB 231 TWLFRYPPLTISQLEIEFRRLRTVKGAEIVGCGDGLGDNFRPTQPLSDRVIRAAF 288

Search completed: March 5, 2004, 09:05:30
Job time: 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2004, 09:00:45 ; Search time 46 Seconds
(without alignments)
3148.323 Million cell updates/sec

Title: US-09-967-237-2
Perfect score: 2424
Sequence: 1 MAFICPSPMLPLIPAPAPG.....RGTGKGVSYPAAVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP virus: *
16: SP bacteriophage: *
17: SP archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	515	21.2	320	6 Q865C0	Q865C0 canis fami
2	493	20.3	344	11 Q8K2T1	Q8K2T1 mus musculu
3	465.5	19.2	317	11 Q7TN69	Q7TN69 mus musculu
4	465.5	19.2	325	11 Q80YB7	Q80YB7 mus musculu
5	430.5	17.8	312	13 Q7SYW3	Q7SYW3 xenopus lae
6	409	16.9	264	11 Q81IX4	Q81IX4 mus musculu
7	397	16.4	260	13 Q7ZK6	Q7ZK6 oncorhynch
8	396.5	16.4	306	13 Q7ZUR2	Q7ZUR2 brachydant
9	395	16.3	260	6 Q865Y7	Q865Y7 bos taurus
10	380.5	15.7	260	11 Q7TPP1	Q7TPP1 mus musculu
11	379.5	15.7	260	13 Q7ZTU6	Q7ZTU6 xenopus lae
12	379.5	15.7	261	13 Q8JG56	Q8JG56 lepidosteus
13	377.5	15.6	260	13 Q8AVG8	Q8AVG8 xenopus lae
14	375	15.5	243	5 Q8MPH8	Q8MPH8 riftia pach
15	368.5	15.2	251	11 Q81IE8	Q81IE8 mus musculu
16	360.5	14.9	261	5 Q9XZG6	Q9XZG6 anthopleura

17	354.5	14.6	311	5 Q9VB76	Q9VB76 drosophila
18	349.5	14.4	327	5 Q9W316	Q9W316 drosophila
19	341	14.1	208	4 Q8ETU0	Q8ETU0 homo sapien
20	337.5	13.9	298	5 Q8WRV4	Q8WRV4 aedes aegypt
21	335	13.8	305	13 Q7SZR2	Q7SZR2 brachydant
22	330	13.6	335	5 Q9VTU8	Q9VTU8 drosophila
23	325	13.4	270	5 Q9V396	Q9V396 drosophila
24	325	13.4	717	11 Q8CINO	Q8CINO ratu
25	325	13.4	1168	11 Q8CIN1	Q8CIN1 ratu
26	325	13.4	1397	11 Q8CIN2	Q8CIN2 ratu
27	325	13.4	1426	11 Q8CIN3	Q8CIN3 ratu
28	314.5	13.0	1576	13 Q91908	Q91908 xenopus lae
29	314.5	13.0	2271	13 Q91909	Q91909 xenopus lae
30	308.5	12.7	429	11 Q8C4W8	Q8C4W8 mus musculu
31	306.5	12.6	275	16 Q8P3B8	Q8P3B8 xanthomora
32	306.5	12.6	192	11 Q9DC73	Q9DC73 mus musculu
33	305.5	12.6	1005	4 Q76043	Q76043 homo sapien
34	302	12.5	271	16 Q8PPA4	Q8PPA4 xanthomora
35	301	12.4	306	11 Q83839	Q83839 mus sp. rec
36	298.5	12.3	240	16 Q8D412	Q8D412 vldro vuln
37	297.5	12.3	247	16 Q7VHP3	Q7VHP3 helicobacte
38	297.5	12.3	1612	11 Q9WDU8	Q9WDU8 mus musculu
39	290	12.0	255	2 Q9XD8	Q9XD8 rhodopseudo
40	288.5	11.9	319	5 Q02637	Q02637 caenorhabdi
41	286.5	11.8	250	16 Q989K6	Q989K6 rhizobium
42	285.5	11.8	235	5 Q9V9Y8	Q9V9Y8 drosophila
43	281.5	11.6	250	2 Q8KJME	Q8KJME rhizobium
44	281.5	11.6	597	11 Q81DU6	Q81DU6 mus musculu
45	279	11.5	277	10 Q9IYE3	Q9IYE3 arabidopsis

ALIGNMENTS

RESULT 1

Q865C0 PRELIMINARY; PRT; 320 AA.
AC Q865C0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Carbonic anhydrase VI precursor.
GN CAC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RA Murakami M.
RT "Canine carbonic anhydrase VI (CAC), mRNA."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB080972; BAC65098.1; -
DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR InterPro: IPR001148; Euk-Coanhd.
DR pfam: PF00194; carb anhydrase; 1.
DR PRODOM: PD000865; Euk-Coanhd; 1.
DR PROSITE: PS00162; Euk_CO2_ANGHYDRASE; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 320 AA; 36704 MW; 6E00A8D08A0D23F CRC64;

Query Match 21.2% Score 515; DB 6; Length 320;
Best Local Similarity 41.6% Pred. No. 2.4e-30;
Matches 107; Conservative 43; Mismatches 95; Indels 12; Gaps 4;

Qy 137 DQSHMYGGDPMPVRVSPACAGRFQSPVDIRPQLAFCAPLRPELILFQQLPPLPRRLR 156
Db 31 DQVH-----WPREYPTGGTROPSPIDLRKRVQVPSLKALKLGYRI-QVGEPPMI 81

```

QY 197 NNGSHVOLTLPPLGEMALGPGREYRALQHLHWGAAGR--PGSEHTVEGHPFAEIHVH 254
Db 82 NNGHTVQISLPPFTMRMASDGTETVIAQMHFHMGASSSEISGSEHTIDIRFVAEIHVH 141
QY 255 LSTAFARVDEALGPRGGLAVLAFL-EGPEENSAVEQLSLREIEAEGSTQVPGDI 313
Db 142 YNSKYKSYDIAQHPEPDLAVLALVKEDEGENTYSNISHLNIRPGQSTVLSGLDI 201
QY 314 SALLPSDFSRFYQYEGSLTTPPCAGVITVFNQVWLSAKQLHTLSDTLWGPQDSRLQL 373
Db 202 EDMLPENTHHYHYRSLTTPPCENTVFWLVHVVHVLSSIQTKLNSILDHQKTLHS 261
QY 374 NFRATQPLNGRVIEASG 390
Db 262 DYRRIQPLNGRVIESNF 278

RESULT 2
Q8K2J1 PRELIMINARY; PRT; 344 AA.
AC Q8K2J1;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN 2310047E01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031385; AAH31385.1; -
DR MGD; MGI:1923709; 2310047E01RIK.
DR GO; GO:0004089; F-carbonate dehydratase activity; IEA.
DR GO; GO:0008270; Zinc ion binding; IEA.
DR GO; GO:0006730; Piome-carbon compound metabolism; IEA.
DR InterPro; IPR001148; Euk_Coand.
DR Pfam; PF00194; Carb_anhydase; 1.
DR ProDom; PD000865; Euk_Coand; 1.
DR PROSITE; PS00162; Euk_CO2_ANHYDRASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 344 AA; 38724 MW; 12DCBD6C3E45D75A CRC64;

Query Match 20.3%; Score 493; DB 11; Length 344;
Best Local Similarity 35.5%; Pred. No. 1,2e-28;
Matches 115; Conservative 50; Mismatches 131; Indels 28; Gaps 7;

```

```

RESULT 3
Q7TNG9 PRELIMINARY; PRT; 317 AA.
AC Q7TNG9;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22386257; PubMed=12477532;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derse J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shennan C.M.; Schuler G.D.;
RA Alyschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Heile F.;
RA Diachenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Schetz T.E.;
RA Brownstein M.J.; Udén T.B.; Toshiki S.; Carrinci P.; Prange C.;
RA Raha S.S.; Loggellano N.A.; Peters G.J.; Adamson R.D.; Mullany S.J.;
RA Bosak S.A.; McKernan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA Villalón D.K.; Wuzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA Fahy J.; Helton E.; Kettelman M.; Madan A.; Rodrigues S.; Sanchez A.;
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Butterfield Y.S.;
RA Krzywinski M.I.; Skalska U.; Smalhus D.E.; Schnerch A.; Schein U.E.;
RA Jones S.J.; Warr M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055437; AAH55437.1; -
KW Hypothetical protein.
SQ SEQUENCE 317 AA; 36464 MW; 6A933F9503E58F0 CRC64;

Query Match 19.2%; Score 465.5; DB 11; Length 317;
Best Local Similarity 37.4%; Pred. No. 1,2e-26;
Matches 98; Conservative 44; Mismatches 107; Indels 13; Gaps 5;

```

```

Q80YB7 ID Q80YB7 PRELIMINARY; PRT; 325 AA.
AC Q80YB7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to carbonic anhydrase 6 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=salivary gland;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049973; AAH49973.1; -
DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR InterPro; IPR001148; Euk_COandh.
DR Pfam; PF00194; carb_anhydrase; 1.
DR Prodom; PD000865; Euk_COandh; 1.
FT NON TER 1
SQ SEQUENCE 325 AA; 37235 MW; 58EB4DA1F02FC121 CRC64;

Query Match 19.2%; Score 465.5; DB 11; Length 325;
Best Local Similarity 37.4%; Pred. No. 1.2e-26;
Matches 98; Conservative 44; Mismatches 107; Indels 13; Gaps 5;

QY 139 SHWRVYGD-----PMPRVSPACAGRFQSPVDIRPOLAFCGALRPLELGFQCPPLPEL 193
DB 27 SDMSYSGDDGSGESQMSQSEYPSGSGERPSIDVKTKEWMPSLKPSLVNVEKENI-EF 85
QY 194 RLNNNGSVQTLPLGLEMALGPGEYRALQHLHWGAAGS-----PGSEHTVEGHRPPAE 249
DB 86 TWNNGHVTSIDLPSMSYLETSDGTEFISKAFHFWG--GSDWELSGSEHTIDGIRIME 143
QY 250 IHVHVLSTAFARVDALGRPGSLAVLA--FLEEGEENSAYEQLSRLBEIAEGSTQV 308
DB 144 AHFVFNKRYGYENAKQKQGLAVLAFKIDEVAENTYSDIISAKIIEKGETTTL 203
QY 309 PGLDISALLPDSFSRYFOYEGSLTTPCAQGVITVFNQVTMLSAKQHLTSLDMLPGCD 368
DB 204 KDTTIRNLPLKDVHHYTYTPGSLTTPCTENVQFVLRDKYTLKSAQVYTIENSVDHNN 263
QY 369 SRLQINFRATQPLNGRVIEASR 390
DB 264 NTLQNGYRSTQPNHRYVEANF 285

RESULT 5
Q7SYW3 ID Q7SYW3 PRELIMINARY; PRT; 312 AA.
AC Q7SYW3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RL MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative."
RL Dev. Dyn. 225:384-391(2002).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Krauss R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Blatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapichon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.D., Ueda T.B., Toshimuki S., Caxiniol P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Guarnatone P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054242; AAH54242.1; -
KW Hypothetical protein.
SQ SEQUENCE 312 AA; 34942 MW; DBC9629136044CF1 CRC64;

Query Match 17.8%; Score 430.5; DB 13; Length 312;
Best Local Similarity 35.4%; Pred. No. 4.6e-24;
Matches 99; Conservative 49; Mismatches 107; Indels 25; Gaps 7;

QY 138 QSHWRVYG-----GDPMPRVSPACAGRFQSPVDIRPOLAFCGALRPLELGFQCP 188
DB 27 EGHWCYELQACDQDCCPFRFNNANMFCGKEQSPINILTKKAVFWESLEKPLKSYGV- 85
QY 189 PLPELRLNNNGSVQTLPLGLEMALGP--GREYRALQHLHWGAAGRPGSEHTVEGHRPP 247
DB 86 SIRQLNTNNGHSAQVLPFGIETLSGGSLGCTDAIQFHHWSEEPFSGSEHTIDDEKIP 145
QY 248 AEIHVHVLSTAFARVDALGRPG3-----LAVLAFLSEGEENSAYEQLSRLBEIAE 302
DB 146 MELHIVH-----RRKTKAKADTGATGSRDLAVLGFEYESTNTNDYERPLNSLAI 199
QY 303 GSETOVPGDISALLP--SDFSRYFOYEGSLTTPCAQGVITVFNQVTMLSAKQHLTSL 360
DB 200 GAVYKPF-SANITKILPKEKELKYIRNGSLTTPCNETVYTWLTFITTKLSQQQURAPY 258
QY 361 DTLNGPQDSRLQINFRATQPLNGRVIEASFPAGVDSPPRA 400
DB 259 NSLFTENORMVENFRVQRLGDRVYIISQQLISSPFS 298

RESULT 6
Q81IX4 ID Q81IX4 PRELIMINARY; PRT; 264 AA.
AC Q81IX4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Carbonic anhydrase 7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RL Chen Y., Huang C.-H.;

```

"Molecular identification of carbonic anhydrases (CA) and CA-related

(CAR) genes";

Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: A1075021; AAL78169.1; "

DR GO: GO:0004089; F:carbonate dehydratase activity; IEA.

DR GO: GO:0008270; F:zinc ion binding; IEA.

DR GO: GO:0006730; P:one-carbon compound metabolism; IEA.

DR InterPro: IPR001148; Euk Coanhd.

DR Pfam: PF00194; carb anhydrase; 1.

DR Prodom: PD000665; Euk_Coanhd; 1.

DR PROSITE: PS00162; Euk_CO2_ANGHYDRASE; 1.

SEQUENCE 264 AA; 29915 MW; B58E0520CB840FA5 CRC64;

Query Match 16.9%; Score 409; DB 11; Length 264;

Best Local Similarity 37.5%; Pred. No. 1,5e-22;

Matches 101; Conservative 41; Mismatches 99; Indels 28; Gaps 11;

DB 141 WRYG--GDPWPRVSPACGRFQSPVDIRPOLAFCALPRLLELGFQPLPELRN 197

DB 7 WGYGDDGDSNMWKLPIAOGDRSPINISSQAVYSPSLQPLLF--YKQWLSITN 63

DB 198 NGSVQL-----TLPRG--LEWALPGREYVALQHLHWGAAGRGSEHTVEGHRFP 247

DB 64 NGSVQVDNDSDDTVTVSGPLF--GP---YRLKQLHFHWGKKRMGSSHTVYDGKSP 117

DB 248 AEIHYVHLST-AFAFVDEALGPGGLAVLAFLLEGPENSAYBQLLSRLLEEIAEGSEET 306

DB 118 SELHLVHWNAKKYSTFGAALAPDGLAVGVFLETG-DEHPSMRLIDALYVRFKDTX 176

DB 307 QVPGDLISALPDSRSRYFQEGSLTTPCAQGYIMVFNQTVWLSKQHTLSDTLM-G 365

DB 177 QSCFNPKCLPFTS-RHYWYTPGSLITPLSESTWVILKRPISRQWEKFRSLFTS 235

DB 366 PGDSRLQ--NFRATQPINGRVTEASFP 392

DB 236 EDDERIHWDNFRPQPLKGRVYKASFOA 264

RESULT 7

Q772K6 PRELIMINARY; PRT; 260 AA.

AC Q772K6; 01-OCT-2003 (TRENBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Erythrocyte carbonic anhydrase (Salmo gairdneri).

OC Oncorhynchus myxiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

NCBI_TaxID=8022;

SEQUENCE FROM N.A.

RA Esbaugh A.J., Lund S.G., Tyfts B.L.;

RT "The Evolution of Red Blood Cell Carbonic Anhydrase: Properties of RBC

CA in a Teleost Fish, Oncorhynchus myxiss."

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY307082; AAP73748.1; "

SEQUENCE 260 AA; 28613 MW; SCC928B72B922A52 CRC64;

Query Match

Best Local Similarity 36.1%; Pred. No. 1,1e-21;

Matches 95; Conservative 38; Mismatches 114; Indels 16; Gaps 8;

DB 139 SH-WRYG--GDPWPRVSPACGRFQSPVDIRPOLAFCALPRLLELGFQPLPELR 194

DB 2 SHWGYGSPDGPEKWCQGFVANGPQSPIDITPGQTSYSTLKPLK---KYDPSNSTD 58

DB 195 LANNGHVQLTLPGLMALGPRG---YRAQLHLMGAARPGSEHTVEGHRPPEAI 250

DB 59 ILNNGHSFQGVFVDDVSSLTGTGPIGTIRLQCFHFWGASDSDRSEHTVNGIKFPCEL 118

DB 251 HVVHLSTAFARVDEALGRPGGLAVLAFLLEGPENSAYBQLLSRLLEEIAEGSEETQVVG 310

DB 119 HLVMNNTKYPSPFGAASEPDGLAVGVFLKIG-AANRLQKVDALCAIKSKQKTFPSN 177

DB 311 LDISALPDSRSRYFQEGSLTTPCAQGYIMVFNQTVWLSKQHTLSDTLM-GPGDS 365

DB 178 FPAKLLLCSD-DYTDGSLTTPPLESTWVILKRPISVSPQNGKFRSLMFGDGEA 236

DB 370 R--LQINFRATQPINGRVTEASF 390

DB 237 PCMODNYRPPQPLKGRKVRSSF 259

Query Match 16.9%; Score 396.5; DB 13; Length 306;

Best Local Similarity 37.3%; Pred. No. 1,5e-21;

Matches 98; Conservative 35; Mismatches 113; Indels 17; Gaps 8;

DB 140 HWRYG--GDPWPRVSPACGRFQSPVDIRPOLAFCALPRLLELGFQPLPELR 196

DB 48 HMGVGEDNGSAMHKKYPIAEGNRQSPIDIVSEAVFPAKSLPIAL---SYNNCTSLIS 104

DB 197 NNGSVQLTLPGLMAL--GP-GREYVALQHLHWGAAGRGSEHTVEGHRPPEAIHV 252

DB 105 NNGSVVVEVDTERSVITGPLENMYRLKQFHFHWSKGGCGGSEHTVAGKTFVSELHL 164

DB 253 VHL-STAFARVDEALGRPGGLAVLAFLLEGPENSAYBQLLSRLLEEIAEGSEETQVPG 311

DB 165 VHWNAKKYKFSZAAPAAPDGLAVLGFLETG-DEHRAHQIDALYVRFKGSIAEFKGF 223

DB 312 DISALPDSRSRYFQEGSLTTPCAQGYIMVFNQTVWLSKQHTLSDTLMGCD--- 368

DB 224 NPKCLPNSL-EWYTPGSLITPPLYESVTWVILKEPIVSEKQKFRSLMFGDEED 262

DB 369 -SRLOINFRATQPINGRVTEASF 390

DB 283 RNEMENNYRPPQPLKGRVYKASFOA 305

RESULT 9

Q865Y7 PRELIMINARY; PRT; 260 AA.

AC Q865Y7; 01-JUN-2003 (TRENBLrel. 24, Created)

DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Carbonic anhydrase II (BC 4.2.1.1).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA Daigle R., Castro I., Desrochers M., Charest P.-M.,
 RT "Full length cDNA of Bovine Carbonic Anhydrase II,"
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY240020; AAC85140.1; -
 DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
 DR InterPro; IPR001148; Euk_Coand.
 DR Pfam; PF00194; carb_anhydrase; 1.
 DR ProDom; PD000865; Euk_Coand; 1.
 DR PROSITE; PS00162; Euk_CO2_ANGHYDRASE; 1.
 DR Lyase.
 KW SEQUENCE 260 AA; 29114 MW; A89412C3FFD19X7 CRC64;
 SQ
 Query Match 16.3%; Score 395; DB 6; Length 260;
 Best Local Similarity 35.9%; Pred. No. 1.6e-21;
 Matches 94; Conservative 42; Mismatches 110; Indels 16; Gaps 8;
 QY 140 HMYRG---GDPPMPRVSPACAGRFQSPVDIRPOLAFCPALPRLLELGFQLPPLRLR 196
 DB 4 HMGSKKNGPENNKKDPFIANGRQSPVIDITRAVHDPALQPL-LISYD--KAASKSIV 60
 QY 197 NNGHSVOLTLPPGLEMAL---GP-GREYRALQLHLWGAGRGSEHTVEGHRPAEIHV 252
 DB 61 NNGHSFVNERDSDQNAVALKGPPLSDSYLLIGFHHMGSSDDGSEHTVNNKKYAAELHL 120
 QY 253 VHLSTARVDEALGRGGLAVLAFLBEGPENSAYEQLSRLEIAEGSETGVPGID 312
 DB 121 VHMNTKYGDGTAAQDPGLAVGVFLKVG-DANPPLQVLAALSIKTKGKSTDPFPFD 179
 QY 313 ISALLPSDFSRFYQEGSLTTPPCAGVITVFNQVMTLSAKOL---HTLSDTLWGPGDS 369
 DB 180 PGLSLPVLND-YMTYGSLLTTPPLESVMTIVLKEPISVSSQMLKFTLNINAGDEFL 238
 QY 370 RLQALNFRATQPLNGRYTEASFP 391
 DB 239 LMLANMRPAQPLKNRQYR-GFP 259
 RESULT 10
 Q7TPE1 PRELIMINARY; PRT; 260 AA.
 ID Q7TPE1
 AC Q7TPE1
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Similar to carbonic anhydrase II.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strauberg R.,
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC041213; AAH41213.1; -
 DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
 DR InterPro; IPR001148; Euk_Coand.
 DR Pfam; PF00194; carb_anhydrase; 1.
 DR ProDom; PD000865; Euk_Coand; 1.
 DR PROSITE; PS00162; Euk_CO2_ANGHYDRASE; 1.
 DR Lyase.
 KW SEQUENCE 260 AA; 28806 MW; ACA02D0D0D818F2C CRC64;
 SQ

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Kodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smalish D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences,"
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RA Strauberg R.,
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC055291; AAH55291.1; -
 DR Hypothetical protein.
 KW SEQUENCE 260 AA; 29041 MW; 43F81AC12238C5E4 CRC64;
 SQ
 Query Match 15.7%; Score 380.5; DB 11; Length 260;
 Best Local Similarity 35.6%; Pred. No. 1.9e-20;
 Matches 93; Conservative 41; Mismatches 112; Indels 15; Gaps 8;
 QY 140 HMYRG---GDPPMPRVSPACAGRFQSPVDIRPOLAFCPALPRLLELGFQLPPLRLR 196
 DB 4 HMGSKKNGPENNKKDPFIANGRQSPVIDITRAVHDPALQPL-LISYD--KAASKSIV 60
 QY 197 NNGHSVOLTLPPGLEMAL---GP-GREYRALQLHLWGAGRGSEHTVEGHRPAEIHV 252
 DB 61 NNGHSFVNERDSDQNAVALKGPPLSDSYLLIGFHHMGSSDDGSEHTVNNKKYAAELHL 120
 QY 253 VHLSTARVDEALGRGGLAVLAFLBEGPENSAYEQLSRLEIAEGSETGVPGID 312
 DB 121 VHMNTKYGDGTAAQDPGLAVGVFLKVG-ASQGLQVLAALSIKTKGKRAAFANFD 179
 QY 313 ISALLPSDFSRFYQEGSLTTPPCAGVITVFNQVMTLSAKOL---HTLSDTLWGPGDS 369
 DB 180 PGLSLPVLND-YMTYGSLLTTPPLESVMTIVLKEPISVSSQMLKFTLNINAGDEFL 238
 QY 370 RLQALNFRATQPLNGRYTEASFP 390
 DB 239 AMVDMRPAQPLKNRQYR-GFP 259
 RESULT 11
 Q7ZYU6 PRELIMINARY; PRT; 260 AA.
 ID Q7ZYU6
 AC Q7ZYU6
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Similar to carbonic anhydrase II.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strauberg R.,
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC041213; AAH41213.1; -
 DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
 DR InterPro; IPR001148; Euk_Coand.
 DR Pfam; PF00194; carb_anhydrase; 1.
 DR ProDom; PD000865; Euk_Coand; 1.
 DR PROSITE; PS00162; Euk_CO2_ANGHYDRASE; 1.
 DR Lyase.
 KW SEQUENCE 260 AA; 28806 MW; ACA02D0D0D818F2C CRC64;
 SQ

Query Match 15.7%; Score 379.5; DB 13; Length 260;
 Best Local Similarity 34.6%; Pred. No. 2.3e-20;
 Matches 90; Conservative 43; Mismatches 112; Indels 15; Gaps 7;

QY 141 WRYG---GDPWPVSPACAGRFQSPVDIRPOLAFCPALPRLLELGFOLPPLRLRN 197
 DB 5 WGYGPDNDPSTWHAFLPANGEXSPINIVPAKAKDQHLKPSI---KYDSTAKVILN 61

QY 198 NGHGVOLTPPLGLEMALGPG---REYRALQHLHMGAGRGSEHTVGHFRPPIHVV 253
 DB 62 NGHAFVNEFPDDSENRSVLSGALSEPRKQHFHFGSCGEGSEHTVGVKCEALHLV 121

QY 254 HUSTAFARVDEALGRPGGLAVLAFLIEEGPEENSAYEQILSRLEETAEBSSETPVPGLDI 313
 DB 122 HMTKYGSMMAVAGHCGGLAVGVFLKVG-EAHPGLQKVLDAKLIPKGNANHSDDP 180

QY 314 SALLPDSRFRFYEGSLTTPPCAQGVITVFNQVWLSAKQHLTSLDTLW-GPGDSRLQ 372
 DB 181 SVLLPNSLD-FWTKYGLTTPPLQLCCVLMHVKLEPIAVSKQQLSRLFFVAEGDTPCS 239

QY 373 L-NFRATOPLNGRVIIEASF 390
 DB 240 MVDNFRPQPLKGRDVRASF 259

RESULT 12
 ID Q8UG56 PRELIMINARY; PRT; 261 AA.
 AC Q8UG56;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Erythrocye carbonic anhydrase.
 OS Lepisosteus osseus (Long-nosed gar).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
 OC Lepisosteus.
 CX NCBI_Taxid=34771;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lund S.G., Dymont P., Gervais M.R., Moyes C.D., Tufts B.L.;
 RT "Characterization of erythrocye carbonic anhydrase in an ancient
 fish, the longnose gar (Lepisosteus osseus).";
 RL J. Comp. Physiol. B, Biochem. Syst. Environ. Physiol. 0:0-0(2002).
 DR EMBL: AY125007; AAM94169.1; -;
 DR GO: GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR GO: GO:0006730; F:zinc ion binding; IEA.
 DR InterPro: IPR001148; Euk Coand.
 DR Pfam: PF00194; carb anhydrase; 1.
 DR Prodom: PD000865; Euk Coand; 1.
 DR PROSITE: PS00162; Euk CO2 ANHYDRASE; 1.
 SQ SEQUENCE 261 AA; 26700 MW; 55BDBF4A08D9F54B CRC64;

Query Match 15.7%; Score 379.5; DB 13; Length 261;
 Best Local Similarity 36.0%; Pred. No. 2.3e-20;
 Matches 95; Conservative 40; Mismatches 112; Indels 17; Gaps 10;

QY 139 SH-WRY---GDPWPVSPACAGRFQSPVDIRPOLAFCPALPRLLELGFOLPPLRLRN 194
 DB 2 SHSWGYAANNPDKHKEKPIAGRQSPIDIVSQAQCHDPLKRLIV---YDPSKSG 58

QY 195 LKRNHGVOLTPPLGLEMAL---GP-GREYRALQHLHMGAGRGSEHTVGHFRPPIHVV 250
 DB 59 IINNGHSFQVVDPAENDSDSLQGGPISGVYRLQFHFHMGASDRGSEHTVGVKAAEL 118

QY 251 HVTALSTA-FARVDEALGRPGGLAVLAFLIEEGPEENSAYEQILSRLEETAEBSSETPV 309
 DB 119 HLYHNMAGKYASFGDAARAPDGLAVGVFLKIGASPR-LQKVIDALDAIKTKGQKPPF 177

QY 310 GUDISALPDSRFRFYEGSLTTPPCAQGVITVFNQVWLSAKQHLTSLDTLW-GPGD 368
 DB 178 QPDKILLPSSLD-FWTKYGLTTPPLLESTVTLVLEKEPIVSSQVAKFRSLFFVAEG 236

QY 369 SRLQL--NFRATOPLNGRVIIEASF 390
 DB 237 TACMVNDYRPPQPLKGRKVRASF 260

RESULT 13
 ID Q8AVG8 PRELIMINARY; PRT; 260 AA.
 AC Q8AVG8;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to carbonic anhydrase II.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 CX NCBI_Taxid=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC042287; AAH42287.1; -;
 DR GO: GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO: GO:0008270; F:zinc ion binding; IEA.
 DR GO: GO:0006730; F:zinc ion binding; IEA.
 DR InterPro: IPR001148; Euk Coand.
 DR Pfam: PF00194; carb anhydrase; 1.
 DR Prodom: PD000865; Euk Coand; 1.
 DR PROSITE: PS00162; Euk CO2 ANHYDRASE; 1.
 SQ SEQUENCE 260 AA; 25004 MW; 2C7BSC0772C0C2DA CRC64;

Query Match 15.6%; Score 377.5; DB 13; Length 260;
 Best Local Similarity 35.0%; Pred. No. 3.2e-20;
 Matches 91; Conservative 41; Mismatches 113; Indels 15; Gaps 7;

QY 141 WRYG---GDPWPVSPACAGRFQSPVDIRPOLAFCPALPRLLELGFOLPPLRLRN 197
 DB 5 WGYGPDNDPSTWHAFLPANGEXSPINIVTAEXKHDKLPISI---KYDSTAKVILN 61

QY 198 NGHGVOLTPPLGLEMALGPG---REYRALQHLHMGAGRGSEHTVGHFRPPIHVV 253
 DB 62 NGHAFVNEFPDDSENRSVLTGALTEPRKQHFHFGSCGEGSEHTVGVKCEALHLV 121

QY 254 HUSTAFARVDEALGRPGGLAVLAFLIEEGPEENSAYEQILSRLEETAEBSSETPVGLDI 313
 DB 122 HMTKYGSMMAVAGHCGGLAVGVFLKVG-EAHPGLQKVLDAKLIPKGNANHSDDP 180

QY 314 SALLPDSRFRFYEGSLTTPPCAQGVITVFNQVWLSAKQHLTSLDTLW-GPGDSRLQ 372
 DB 181 SVLLPNSLD-FWTKYGLTTPPLQLCCVLMHVKLEPIAVSKQQLSRLFFVAEGDTPCS 239

QY 373 L-NFRATOPLNGRVIIEASF 390
 DB 240 MVDNFRPQPLKGRHTRASF 259

RESULT 14
 ID Q8MPH8 PRELIMINARY; PRT; 243 AA.
 AC Q8MPH8;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Carbonic anhydrase (EC 4.2.1.1).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Mammalia; Eumammalia; Eutheria; Rodentia;
 OC Rattidae; Rattidae; Rattidae.
 CX NCBI_Taxid=6426;
 RN [1]

RP SEQUENCE FROM N.A.
 De Cian M.C., Bailly X., Boulben S., Strub J.M., Von Dorsallaer A.,
 Lallier F.H.;
 "An insight into molecular and biochemical characteristics of carbonic
 anhydrases from Riffia pachyptila, a symbiotic invertebrate living
 under extreme conditions."
 RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AJ439711; CAD29128.1; -
 DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
 DR InterPro: IPR001148; Euk Coahnd.
 DR Pfam: PF00194; carb_anhydrase; 1.
 DR Prodom: PD000865; Euk_Coahnd; 1.
 DR Prosite: PS00162; Euk_CO2_ANHYDRASE; 1.
 KM Lyase.
 SQ SEQUENCE 243 AA; 25661 MW; AAA53B407A1P4645 CRC64;

Query Match 15.5%; Score 375; DB 5; Length 243;
 Best Local Similarity 35.3%; Pred. No. 4.6e-20;
 Matches 90; Conservative 36; Mismatches 109; Indels 20; Gaps 8;

QY 141 WRYGDDP--WPRVSPACGRFGSPVDIRPOLAFCPLRPLELGFQLPPLPELRNN 198
 DB 4 WDYEANGPATWAKSFPLAGKKQSPIDIPASV---SKKSTALVASYNPAASNTLTNT 59
 QY 199 GHSVQLTPGLEMALGP-GREYRALQHLHMGAGRPGSEHTVESHGRRPAETHVHLST 257
 DB 60 GLSFQVSVYDGLT--SGGPLNGEYKASFFHFWKTSABGSEHTVAGKAYAAEAHIVHNA 117
 QY 258 A-FARVDEALGRPGGLAVLAFLFEGPEENSAYEQLLSRLEETAEESGTQVP-GLDISA 315
 DB 118 AKYASFDQAVADGGLAVLAFIOPG-ATNAGVQKIIDLPSVPTGDTATIPGSPDVAC 176
 QY 316 LLPSDSRFRFOYEGSLTTPPCAGCVITVFNQVMTLSAQQLHTLSDTLWPGDSRLQINF 375
 DB 177 LLPGDQSKYWYPPSLTTPPCFESVTWIKDPIQCEQALALRKI-----TGCNF 228
 QY 376 RATQPLNGRVIETASF 390
 DB 229 RPTIGLCGROYSSSF 243

RESULT 15
 Q811E8 PRELIMINARY; PRT; 251 AA.

ID Q811E8
 AC Q811E8;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to carbonic anhydrase 6 (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=salivary gland;
 RA Strauberg R.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC046495; AAH46495.1; -
 DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
 DR InterPro: IPR001148; Euk_Coahnd.
 DR Pfam: PF00194; carb_anhydrase; 1.
 DR Prodom: PD000865; Euk_Coahnd; 1.
 DR NON TER 1
 FT NON TER 1
 SQ SEQUENCE 251 AA; 29182 MW; 17F966486FB4261D CRC64;

Query Match 15.2%; Score 368.5; DB 11; Length 251;
 Best Local Similarity 38.2%; Pred. No. 1.4e-19;

Matches 78; Conservative 34; Mismatches 85; Indels 7; Gaps 3;
 QY 192 ELRRNNGHSVQLTPPGLEMALGPGRYRALQLHMGAGR---PGSEHTVESHGRRP 247
 DB 10 EFTTNGHGVISIDLPSSMYLTSDDTEFISKAFFHWG--GRDWELSGSEHTIDGIRSI 67
 QY 248 AEIHVHALSTAFARVDEALGRPGGLAVLAFLFEGPEENSAYEQLLSRLEETAEESGSET 306
 DB 68 MEAFVAFNKEVGTYEAKQKNGGLAVLAFKIDYAEANTYSDIISAKNIEKGETT 127
 QY 307 QVPELDISALLPSDSRFRFOYEGSLTTPPCAGCVITVFNQVMTLSAQQLHTLSDTLWGP 366
 DB 128 TLKDTIRNLLPXDVHHYTYPPSLTTPCTENVQFVLDRKVTLSKAQVVTIENSVMH 187
 QY 367 GDSRLQNLPRATQPLNGRVIETASF 390
 DB 188 NNNTIQNGYRSTQPNHRYVEANF 211

Search completed: March 5, 2004, 09:06:30
 Job time : 48 secs

RESULT 1
138013
p54/58N - human
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-Nov-1999
C:Accession: 138013
R:Pastorek, U.; Pastorekova, S.; Callebaut, I.; Mornon, J.; Zelnik, V.; Opavsky, R.; Zatl
Oncogene 9, 2877-2888, 1994
A:Title: Cloning and characterization of MN, a human tumor-associated protein with a dom
A:Reference number: 138013, MVID:94366734, PMID:8084592
A:Accession: 138013
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-459 <RES>
A:Cross-references: EMBL:X66839; NID:G1000701; P1DN:CAA47315.1; P1D:G1000702
C:Genetics:
A:Gene: MatuMN
C:Superfamily: carbonic anhydrase homology
F:141-390/Domain: carbonic anhydrase homology <CAH>

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 4, 1e-147;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAPLCSPWLPPLIPAPAPGLTVOLLISLLIMPVHPORLPROMEDSFLGGSSGEDDPL	60
DB	1	MAPLCSPWLPPLIPAPAPGLTVOLLISLLIMPVHPORLPROMEDSFLGGSSGEDDPL	60
QY	61	GEEIDLSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP	120
DB	61	GEEIDLSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP	120
QY	121	DPEPQNNARHKGDDQSHMRVGGDPMPVSPACAGRFQSPVDIIRPOLAFCPALRPL	180
DB	121	DPEPQNNARHKGDDQSHMRVGGDPMPVSPACAGRFQSPVDIIRPOLAFCPALRPL	180
QY	181	ELIGFQLPPLPFLRLRNNGHVSQVLTLPGLMALGPGREYRALTQHLHMGAGRGSEHT	240
DB	181	ELIGFQLPPLPFLRLRNNGHVSQVLTLPGLMALGPGREYRALTQHLHMGAGRGSEHT	240
QY	241	VEGHRPPELHVHSTFAVDEALGRPGCLAVLAFLBEGPEENSAVQQLSRLEEA	300
DB	241	VEGHRPPELHVHSTFAVDEALGRPGCLAVLAFLBEGPEENSAVQQLSRLEEA	300
QY	301	EEGSETOVPGDLISALPDSFRYFOYEGSLTTPCAQGVITWTFNOTVMSAKQHTLS	360
DB	301	EEGSETOVPGDLISALPDSFRYFOYEGSLTTPCAQGVITWTFNOTVMSAKQHTLS	360
QY	361	DTLMGPGSRQLNFRALQPLNGRVTEASPPAGTDSPPRAEPVQVNSCLAAGDILALVF	420
DB	361	DTLMGPGSRQLNFRALQPLNGRVTEASPPAGTDSPPRAEPVQVNSCLAAGDILALVF	420
QY	421	GLFAVTSVAFLVOMKRRHGRGTGKGVSRPAEVAETGA	459
DB	421	GLFAVTSVAFLVOMKRRHGRGTGKGVSRPAEVAETGA	459